Unifying the Demand and Supply-Side Theories of Minority Descriptive Representation: A Logical Model

Yuki Atsusaka*

atsusaka@rice.edu

March 18, 2020

Abstract

In this article, I derive a quantitatively predictive logical model of minority candidate emergence. Drawing on previous research, I demonstrate that the probability of minority candidates running for office can be computed as $\mathbf{F}((MC)^{1/2} - 50)$, where M is what I call the racial margin of victory, C is the percentage of minority voters, and $\mathbf{F}$ is some cumulative distribution function. To show the validity of the model, I visualize its quantitative predictions and test them with a novel data set of Louisiana mayoral elections from 1986 to 2016. The logical model can correctly predict 88% of minority candidate emergence and its in-sample and out-of-sample predictive performance is higher than regressions. The proposed model unifies two competing theories of minority candidate emergence and provides a new insight that the district racial composition is only half of the story of minority descriptive representation.

*Yuki Atsusaka is a graduate student in the Department of Political Science at Rice University (https://atsusaka.com).
For helpful comments, I thank Chano Arreguin-Abarca, Matthew Hayes, Melissa Marschall, Kaitlin Senk, Randy Stevenson, Tyler Reny, Liana Reyes-Reardon, M.C. Sunny Wong, Marques Zarate, and the participants of the University of Houston EITM Summer Institute 2019.

1 Introduction

Under what conditions do racial minority candidates enter electoral contests? Understanding the causal mechanisms of minority candidate emergence (and victory) has been of great interest for scholars of minority representation, entailing critical normative and policy implications related to the Voting Rights Act (VRA) and redistricting (Grofman, Handley and Niemi, 1992; Elmendorf, Quinn and Abrajano, 2016). Previous research has provided two competing theories of minority candidate supply (e.g., Shah, 2014; Juenke, 2014; Juenke and Shah, 2015, 2016; Fraga, Juenke and Shah, 2019) and voter demand (e.g., Abosch, Barreto and Woods, 2007; Barreto, Segura and Woods, 2004; Trounstine and Valdini, 2008; Lublin et al., 2009). In contrast, this article shows that both theoretical arguments can be integrated under a unified theory of minority candidate decision-making. The proposed logical model offers a new insight that the size of minority voters – the most important factor in previous research – is only half of the story and minority candidates learn about voter demand from electoral performance of co-ethnic candidates in previous elections.

This article makes several contributions. First, I derive a *quantitatively predictive logical model* of minority candidate emergence drawing on an emerging approach in comparative politics (Taagepera, 2007, 2008; Li and Shugart, 2016; Shugart and Taagepera, 2017; Taagepera and Nemčok, 2019). Building on the literature of political ambition, I demonstrate that the probability of minority candidates running for office can be computed as $\mathbf{F}((MC)^{1/2} - 50)$, where M is what I call the racial margin of victory (the difference in the vote shares obtained by the top minority and top white candidates, see Section 3), C is the percentage of minority voters, and $\mathbf{F}$ is some cumulative distribution function. Next, I illustrate the logical model by visualizing its quantitative predictions and deriving two observable implications as hypotheses: (1) the racial margin of victory in the last elections increases the probability of minority candidate emergence and (2) the degree of such influence is the greatest in districts where minority and white voters are equally distributed. Finally, I test the validity of the logical model with a novel data set of Louisiana mayoral elections from 1986 to 2016. I demonstrate that the logical model has remarkably high predictive performance both in-sample and out-of-sample and it can correctly predict 88% of minority candidate emergence and 94% of minority candidate victory. It is then shown the "logical" model (obtained without looking at any data) dominates linear and logistic regressions ("statistical" models) in predictive performance, while the model prediction also holds in causal analysis.

As elaborated below, the innovation of this study is to provide a parsimonious mathematical model of minority representation which is (1) based on deductive logic – and not on any statistical analysis – and (2) able to offer quantitative predictions about the probability of minority candidate emergence (our quantity of interest); hence the name quantitatively predictive logical models (Taagepera, 2008). The logical model approach tightly aligns the theoretical concepts of interest and their empirical measures as required in good model building (Granato, Lo and Wong, 2010). This is highly advantageous in scientific studies of politics since it can mitigate researchers' degree of freedom in measuring concepts and choosing functional forms of variables in statistical models (along with the impact of sampling variability in data) and thus ease the cumulation of knowledge.

Another advantage of the logical model is that it establishes a logical justification for a particular functional form of relevant variables. For example, the proposed model includes the percentage of minority voters C in a multiplicative form $(MC)^{1/2}$ to represent a theoretical argument that minority candidates take into account for both the racial **m**argin of victory in the last elections (M) and the expected racial margin of victory based on the racial **c**omposition (C) when calculating the probability of winning. In Online Appendix A, I demonstrate that in the first-past-the-post elections, C is algebraically identical to the percentage of minority voters. As I discuss in the appendix and Section 3, the multiplication is the result of taking the "geometric mean" of the two quantities. This is in contrast to many research in this area that includes the size of minority voters as a "to-go" variable without any logical justification beyond the directional hypothesis that it will increase the likelihood of minority candidate emergence.

Theoretically, the novelty of this article is that it theorizes minority candidate emergence as an "interaction" between candidates' strategic calculation and voters' demand. While recent studies focusing on the supply-side of minority representation were successful in bringing candidates in the discussion (Shah, 2014; Juenke, 2014; Juenke and Shah, 2015, 2016; Fraga, Juenke and Shah, 2019), the conflict between the demand and supply-side theories is rather misleading. This article echos several other studies in emphasizing the complementary roles of both sides (Bullock III and Johnson, 1985; Bullock III and Smith, 1990; Canon, 1999; Fraga, 2014) and formulates such argument as a logical model. Consequently, the logical model provides a key to understand a set of cases the conventional approach cannot explain, such as the absence of minority candidates in majority minority districts and the presence of minority candidates in majority white districts (Juenke and Shah, 2016; Shah, 2017).

In what follows, I begin by outlining the two competing theories of minority candidate emergence and clarify their policy and normative implications. I then detail my theoretical argument informally; specify the scope condition; and derive the logical model. After introducing the logical model, I visually describe its several observable implications and draw two hypotheses. Finally, I demonstrate the validity of the logical model by testing its predictive performance and other observable implications. Additional analysis, information, and discussion can be found in Online Appendix that supplements this article.

2 Voter-Demand and Candidate-Supply Theories

To explain the relative lack of minority descriptive representation, scholars have considered two competing theories, which include voter-demand and candidate-supply theories (Shah, 2014; Juenke and Shah, 2015, 2016). The demand-side theory sees the relative dearth of minority politicians as the minority candidate *defeat* problem. Here, minority candidates cannot win electoral contests outside of majority minority districts due to strong opposition from white voters (Abosch, Barreto and Woods, 2007; Lublin et al., 2009; Trounstine and Valdini, 2008). Thus, the relative lack of minority representatives is the result of white voters who do not vote for minority candidates. An important normative and policy implication of the voter demand theory is that creation and maintenance of majority minority districts is *the* solution to minority underrepresentation (Barreto, Segura and Woods, 2004; Bedoya, 2005; Casellas, 2010; Lublin, 1999; Lublin et al., 2009).

In contrast, more recent scholarship based on the supply-side theory describes minority underrepresentation as the minority candidate *retreat* problem. Here, minority candidates do not run for office outside of majority minority districts even though they have good chances of winning once they enter electoral competitions. Therefore, the lack of minority representatives is due to minority potential candidates' miscalculation of the odds of winning (Shah, 2014; Juenke, 2014; Juenke and Shah, 2015, 2016; Fraga, Juenke and Shah, 2019). One critical implication of the candidate supply theory is that minority underrepresentation can be partly solved by providing minority potential candidates a set of high quality information about electoral fortunes and voting behavior of the white electorate in their districts.

While each of the two competing theories may seem plausible, several unsolved problems remain. First, previous research has not explicitly modeled the decision-making of minority potential candidates. Additionally, most research has treated racial bias by white voters as a black box or a fixed phenomenon. In other words, the literature has recognized that white voters are less likely to vote for minority candidates, but it does not illustrate to what extent such racial bias is present in different elections and how minority candidates respond to varying degree of racial bias if any. Moreover, it is unclear what constitutes minority candidates' self-selection into majority minority districts, which appears to be a key part of the candidate supply theory of minority representation. Finally, most research has not integrated different types of districts – or what I call a *racial regime* – and the temporal dimension in the discussion of minority descriptive representation. By the racial regime, I refer to a different type of district defined by the racial composition such as majority minority, racially balanced, and majority white districts, which might influence minority candidates' strategies and calculation of the likelihood of winning. By the temporal dimension, I describe a connection between current elections and past elections in terms of electoral performance of minority candidates relative to their white counterparts, which might also affect how minority candidates perceive their likelihood of winning. While the literature has acknowledged the importance of considering such aspects of minority representation, no single research has incorporated both factors in a coherent framework.[1]

In the next section, I describe my proposed model and illustrate how it unifies both voter-demand and candidate-supply theories of minority descriptive representation.

3 A Unified Model of Minority Candidate Emergence

In this section, I present a logical model of minority candidate emergence, elaborate and visualize observable implications from the model, and then draw two hypotheses that I test in later sections. Before introducing formal arguments, let me informally describe my theoretical argument about when and where minority candidates decide to run for office.

In the proposed model, I attempt to understand the district level minority candidate emergence as the decision making of the most viable minority politicians, whose sole agenda is to win elections where two racial groups (i.e., minority and majority groups) compete each other. Given these assumptions, I theorize that minority candidates decide to run for office when they see a higher probability of winning. Minority candidates then attempt to calculate the prior likelihood of winning in the upcoming elections, but as for

[1] It must be noted that, separately, researchers have already taken up on the issue of racial regime and temporal dimension. For example, Juenke and Shah (2016) examine minority candidate emergence in white, racially mixed, and majority minority districts and report interactive effects of racial regimes, candidate race, and candidate partisanship. Shah (2017) also focuses on three "racial profiles" including majority white, majority minority, and multiracial cities and elaborates different power dynamics between racial groups in different types of cities. On the temporal dimension, Marschall, Ruhil and Shah (2010, 114-15) discuss that past success in minority electoral bits may influence minority candidate victory in later years, although they use the past representation variable as a control variable. Similarly, Shah (2014, 269, 271) considers the impact of any history of black candidacy on the current black minority emergence based on the theoretical insight that the initial hurdle to run for office as minority politicians is always the hardest in minority representation. Shah (2017) also studies how the changes in racial composition would affect minority electoral fortune and candidate supply.

any candidate, they are considered to be bounded rational and thus try to make the most satisfactory choice based on incomplete information. Because it is quite difficult for minority candidates, as for any candidate, to calculate the prior likelihood of winning, they rely on two sources of information, which consist of (1) electoral performance of co-ethnic candidates in the last elections and (2) district racial composition as relevant heuristics. I then claim that the value of information from the last elections increases as districts become more competitive or more racially heterogeneous because the racial makeup is not as informative in racially balance districts as in racially homogeneous districts.

In the rest of this section, I formalize my argument and provide a logical model of minority candidate emergence.

3.1 Quantitatively Predictive Logical Models

To consider when and where minority candidates emerge, I focus on the strategic entry of the most viable minority candidate in each electoral district. To model minority candidate emergence, I adopt an emerging approach in comparative politics to derive a *quantitatively predictive logical model* (Taagepera, 2007, 2008; Li and Shugart, 2016; Shugart and Taagepera, 2017; Taagepera and Nemčok, 2019). One prominent feature of this approach is that it relies on deductive logic to derive a parsimonious mathematical equation (without looking at any data!) which can accurately predict the outcome of interest.[2] This logical modeling is in contrast to a more conventional approach in political science where researchers include a number of independent variables to multivariate regressions and deductively provide a logic to understand why certain coefficients are statistically significant or not *after* running their regressions.

Another feature is that the logical modeling approach has a tighter connection between theoretical concepts and their empirical measures, which enables researchers to provide stronger empirical tests for their theoretical arguments. More importantly, the logical modeling has clearer policy implications because the model can exactly tell *how much* the outcome changes (i.e., quantitative predictions) if we vary the parts of the equation. Therefore, the logical model can offer a considerably helpful information for important practical problems including electoral engineering and constitutional design, which have enormous impacts on representation and democratic process (Li and Shugart, 2016, 23). While such logical modeling has hardly appeared in the literature of minority representation, and political science more broadly, it appears

[2] A canonical example is the Seat Product Model developed by Taagepera (2007), where the effective number of parties with legislative representation (N_s) is predicted by a function of a country's average district magnitude (M) and its assembly size (S) as $N_s = (MS)^{1/6}$.

to be more common in other scientific disciplines (Colomer, 2007). I must emphasize that by definition the following logical model is a parsimonious representation of the reality and does not explain every possible detail in minority representation. The essence of the approach is that even though it only considers two variables, it can correctly predict 88% of minority candidate emergence as shown in Section 5. Thus, readers are asked not to dismiss the argument based on its style, but to judge it on its observable implications and empirical evidence for them.

Below, I start by introducing four assumptions and then describe the logical model of minority candidate entry.

3.2 Scope Condition

Generally speaking, a good model requires a clear scope condition that the model can speak to. Thus, to clarify which kinds of electoral contexts the model accounts for, I limit the scope condition by stating four assumptions.

Assumption 1 (Biracial elections). *Electoral competitions are held over two racial groups, majority and minority, and candidates' race is one of the most prominent factors which affect people's voting behavior based on the strong tendency of co-ethnic voting.*

The first assumption states that I consider classic biracial elections where two racial groups compete each other and voters tend to support co-ethnic candidates.

Assumption 2 (Non-zero pool). *There is always a non-zero number of minority politicians or potential candidates who could run if conditions allowed in each district.*

This assumption excludes the possibility that I do not observe any minority candidate running for office due to the lack of the "supply" of minority potential candidates in the pool.[3]

Assumption 3 (Instrumental candidates).*Minority candidates are short-term instrumental such that their primary goal is to get elected in elections.*

[3] For studies looking at the potential-candidate pool, see Maisel and Stone (1997) and Fox and Lawless (2004).

This assumption rules out the possibility that minority candidates decide to run for office due to non-instrumental reasons such as symbolic reasons in which they seek to run for office to obtain benefits from the action of running itself (e.g., raising voice or selling names for future elections).

Assumption 4 (The most viable candidate).*Whether I see at least one minority candidate or not solely depends on the strategic choice of the most viable minority politician in the candidate pool in the electoral district.*

This assumption enables us to model the binary process of minority candidate emergence, which can be only observed at the district level, as the individual decision making by the most viable minority politician in the district. While I consider the decision-making of minority candidates and their teams such as party stuff and strategists, I use the term minority "candidates" to suppress the complication. The last assumption implies that when and where minority candidates are observed does not depend on the behavior of those candidates whose perceived odds of winning is not the highest among co-ethnic potential candidates. I believe this assumption to be plausible based on Assumption 3 that minority candidates are rational office-seeking actors and that less viable candidates would not decide to enter the electoral competition unless the most viable candidates do so. Thus, along with Assumption 1, this can be considered as a racial version of an "M+1" rule (Reed, 1990; Cox, 1997).[4]

In summary, Assumptions 1-4 claim that the district level minority candidate emergence can be modeled as the decision making of the most viable minority politicians whose sole agenda is to win classical biracial elections. It should be emphasized that my theory does not consider elections with primary and general elections (Stone and Maisel, 2003) and such extension must be explored in future research.

3.3 The Logical Model

Based on these assumptions, I derive the following logical model of minority candidate emergence. The derivation of the model is detailed in Online Appendix A.

[4] An "M+1" rule posits that the number of serious or viable candidates (or parties) in a district tends to be its district magnitude (M) plus one (Reed, 1990; Cox, 1997). In the first-past-the-post elections, where $M = 1$, the rule suggests that the number of serious candidates tend to be 2.

$$\Pr(E) = \widehat{P} = \mathbf{F}\big((MC)^{1/2} - 50\big) \quad (1)$$

The logical model suggests that the probability of minority candidate entry, $\Pr(E)$, is equal to the estimated probability that minority candidates win the electoral contest once they run for office $\widehat{P}$. It further shows that such estimated probability can be computed by two variables M and C through some cumulative distribution function (CDF) **F** such as a standard normal CDF known as a probit function.[5]

M is what I call the *racial* ***m****argin of victory* in the last elections and it quantifies the extent to which minority candidates secured their descriptive representation relative to their white counterparts. More precisely, it is based on the difference between the vote shares that went to the top minority candidate and the top white candidate in the last election at each district: $M = \frac{1}{2}(V^M_{t-1} - V^W_{t-1}) + 50$.[6] Drawing on the concept of the margin of victory, it quantifies the degree of vote shares that need to be modified in order to change the "race" of the winner.

Similarly, C denotes the *expected racial margin of victory* in upcoming elections based on the racial ***c****omposition* of districts. It expresses the racial margin of victory under the assumption that minority and white voters only vote for their co-ethnic candidates. Surprisingly, after some algebra, it turns out that C is equivalent to the percentage of minority voters.[7]

The model then takes an average of the two quantities via a "geometric mean," $(MC)^{1/2}$, to represent the argument that minority candidates take into account *both* information M and C as relevant heuristics to calculate a future probability of winning.[8] I specifically use geometric means since they have attractive mathematical and theoretical properties in logical model building (Taagepera, 2008). Finally, without changing any substantive reasoning, 50 is subtracted from the geometric mean for adjusting the transformation applied to M and C.[9]

To sum up, the logical model provides quantitative predictions about minority candidate emergence with only two terms. Importantly, both terms are based on widely agreed-upon measurements in political science and, thus, theoretical concepts and measurements are tightly aligned in the model as required in good model

[5]While the model can be written as $\Pr(E_{it}) = \widehat{P}_{it} = \mathbf{F}\big((M_{it-1}C_{it})^{1/2} - 50\big)$ to index district by i and time period by t, I suppress such indices to follow the convention in logical model building (Taagepera, 2007, 2008; Shugart and Taagepera, 2017).

[6]Technically, M is an affinely transformed racial margin of victory as $M = \tilde{M} + 50$, where $\tilde{M}$ is the raw margin of victory.

[7]To show this, let me simply state that the raw expected racial margin of victory becomes $\tilde{C}$ =% Minority Voters - 50. After applying the same affine transformation, $C = \tilde{C} + 50$ = % Minority Voters.

[8]A geometric mean of N numbers $(y_1, ..., y_N)$ is defined as $\sqrt[N]{y_1 *, ..., * y_N} = (\prod_{i=1}^{N} y_i)^{1/N}$.

[9]The raw racial margin of victory in last elections $\tilde{M}$ and the raw expected racial margin of victory $\tilde{C}$ are transformed as $M = \tilde{M} + 50$ and $C = \tilde{C} + 50$ in order to avoid the multiplication of negative numbers.

building (Granato, Lo and Wong, 2010).[10] It is also worth emphasizing that the model was derived from a deductive logic and not stemmed from any statistical analysis. Consequently, the model's predictions are highly stable and not susceptible to how researchers measure concepts and what sample or training data they use as in other types of predictive models.

Theoretically, what the logical model indicates is that minority candidate emergence is based on minority candidates' decision-making process as advocated by the supply-side theory, whereas their decision-making is also largely affected by past and potential vote shares as the demand-side theory expects. In other words, the model claims that minority candidates react to past voter demand M to foresee future electoral fortune given the district racial composition C. This way, the logical model unifies the two theoretical perspectives which have been contrasted against each other in previous research. Therefore, it provides a more comprehensive picture of minority candidate emergence, but it does not in a very parsimonious manner. In the next section, I demonstrate that not only does the model make sense on the logical ground, but also it offers a remarkably high predictive power on minority candidate emergence. Before showing the results, I clarify several observable implications from the logical model.

3.4 Observable Implications

To illustrate how the model works, let me first provide a simple numerical example. To contextualize the example, assume that black voters are the minority group of interest and white voters as the majority group of reference. Suppose that I observe that the top black candidate and top white candidate obtained 30% and 50% of vote shares, respectively, in the last election.[11] Suppose also that the district is composed of 60% black voters and 40% white voters. This exemplifies the district where black voters are slight majority but a white candidate won the last election.

With these information, it is easy to show that $M = \frac{1}{2}(30 - 50) + 50 = 40$ and $C = 60$. The model then suggests that the probability of minority candidate entry becomes: $\mathbf{F}((40 * 60)^{1/2} - 50) = \mathbf{F}(48.99 - 50) = \mathbf{F}(-1.01)$. When a probit function is used for $\mathbf{F}$, the probability of black emergence becomes $\Phi_{0,1}(-1.01) = 0.156$ or there is 15% chance that a black candidate appears in an upcoming election in the same district.

[10]While the racial margin of victory is a novel concept that I introduce in this article, it is based on a conventional idea of the margin of victory. For details, see Online Appendix A.

[11]Importantly, these vote shares need not to be summed up to 100% when multiple candidates are on ballots and they obtain some portion of the total ballots.

Figure 1 visualizes the logical model's quantitative predictions under varying conditions. The left panel displays the probability of minority candidate emergence against $MC^{1/2} - 50$ with different CDF for **F**. When the term is negative, the probability is lower than or equal to 0.5, whereas it becomes higher than or equal to 0.5 when the term is positive. Substantively, thus, $MC^{1/2} - 50$ represents a (racial) margin of victory and the shape of **F** controls the degree of uncertainty around the margin of victory. Nevertheless, the overall look of the S-curve does not change greatly regardless of the choice of **F**.

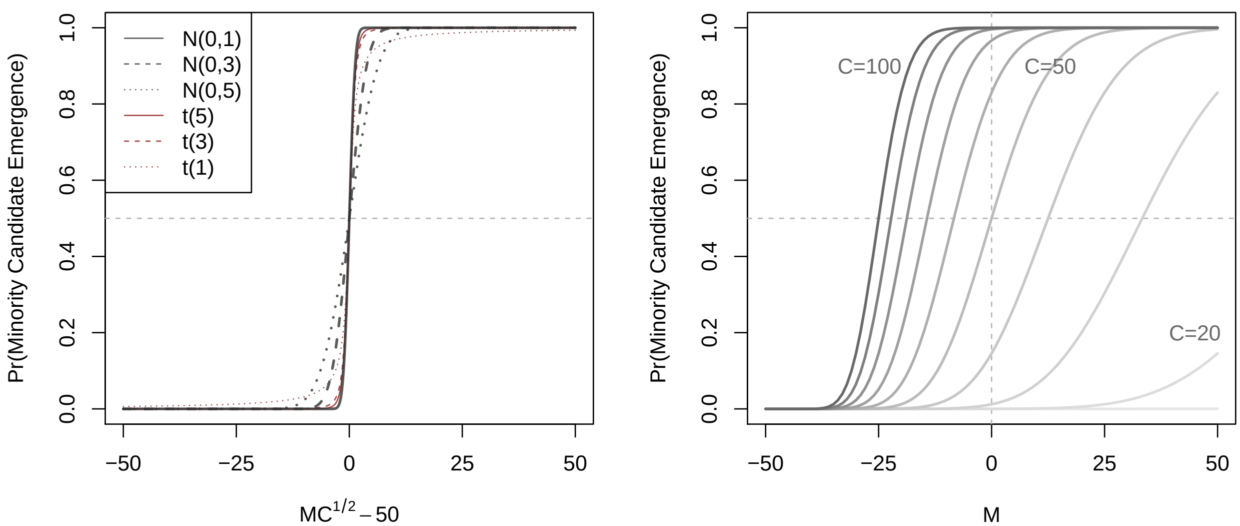

Figure 1: **Quantitative Predictions of the Logical Model**
Note: The left panel shows the probability of minority candidate emergence as a function of $MC^{1/2} - 50$ with various forms of **F**. The right panel visualizes the impact of M (racial margin of victory) on the probability under varying the levels of C (the percentage of minority voters).

At first glance, the two terms in the logical model seem to have equal impacts on the probability of candidate emergence. Somewhat surprisingly, this is not an accurate description of the model. To see why, see the right panel of Figure 1 where I plot the candidate emergence probability against M with varying values of C. It illustrates that (1) increase in M raises the probability, but (2) the degree of such increase depends on the value of C. For example, when $C = 20$ (20% minority districts), the probability of minority candidate emergence is mostly 0 and it starts to marginally increase after M is beyond 25. Theoretically, this means that minority candidate is not likely in the 20% minority districts *unless* black candidates were elected with large margins (did "extremely well") in the last elections.

The opposite case can be seen when $C \approx 100$ (predominantly minority districts). Here, the probability of minority candidate running is mostly 1 and it starts to decline when M is less than -10. Substantively, this indicates that minority candidates tend to be on ballots with certainty *unless* they were defeated by white candidates with great margins (did "extremely poorly") in the last elections. Finally, the influence of M on the probability becomes the greatest when $C = 50$ (50% minority districts). The logic behind this property is that the district racial composition has no information about minority candidates' expected vote shares under Assumptions 1-4, and thus what happened in the last elections become the most informative. In other words, the logical model suggests that minority candidates as decision-makers put more weights on the information from the last election as the district becomes more racially balanced.[12]

The innovation of the logical model is that it can provide "quantitative" predictions (i.e., exact probability in this context) as opposed to qualitative or directional predictions (i.e., the increase or decrease in the probability); hence the name quantitatively predictive logical models (Taagepera, 2008; Shugart and Taagepera, 2017). Finally, I explicitly state the following observable implications from the logical model as testable hypotheses (King, Keohane and Verba, 1994).[13]

Hypothesis 1. *We are more likely to observe minority candidates in districts where the racial margin of victory in the last elections (M) was high.*

Hypothesis 2. *The degree of influence that the racial margin of victory (M) has on minority candidate emergence is greater in racially balanced districts ($C \to 50$) than in majority white ($C \to 0$) and majority minority districts ($C \to 100$).*

4 Data

To validate the model predictions, I construct a novel candidate-level data set of mayoral elections in 313 Louisiana municipalities from 1986 to 2016.[14] Louisiana mayoral elections provide a great test case for the model predictions because they use a unique electoral system called the majority run-off system, which

[12]The asymmetry seen in the right panel of Figure 1 is derived from the fact that the logical model is based on the "geometric" mean (as opposed to arithmetic mean) of the two terms.

[13]Here, I do not explicitly hypothesize that the percentage of minority voters (C) are positively associated with minority candidate emergence. The size of minority voters has been the first and foremost variable that previous studies include in their regressions even though they do so without clear quantitative predictions. The contribution of this study is to provide a logical justification for including the percentage (or proportion) of minority voters in the analysis.

[14]This data set was complied and processed as a part of the Local Elections in America Project (LEAP) (Marschall and Shah, 2013) between 2016 and 2017.

enables us to overcome a potential problem of using general election data to study minority candidate emergence. In general elections, the absence of minority candidates stems from two possibilities: (1) no minority candidate decided to run for office to begin with and (2) minority candidates emerged but were defeated in "primary" elections. It must be emphasized that previous research using general election data does not (as it cannot) empirically differentiate the two potential mechanisms (Juenke, 2014; Juenke and Shah, 2015, 2016; Fraga, Juenke and Shah, 2019). In contrast, in the majority run-off system, all candidates participate in open-primary elections regardless of partisan affiliation, and the candidate with the majority votes becomes the winner (Keele et al., 2017).[15] This enables us to eliminate the second possibility for the absence of minority candidate, and make inferences about the emergence of minority candidates as the direct consequence of minority candidates' decision to run for office.

Louisiana elections also serve as a great benchmark for verifying the logical model since more than 96% of voters are either African American or white (according to the official registration records with self-reported race, see Online Appendix B.5). The data also contains vote totals for all 5297 candidates in 2037 elections, and it is collapsed at the election-level so that each contest becomes the unit of analysis. Moreover, I compiled information about candidates' race based on internet and news article search and thus presents a unique opportunity to examine minority candidacy with more accuracy than other race imputation or inference methods (Shah and Davis, 2017).

As the logical model predicts varying effects of the racial margin of victory conditional upon the district composition (H_2), I employ an organizing principle of the racial regime to examine such effects. Specifically, I subset the entire data of municipalities into majority white, racially balanced, and majority minority (black) districts by the average percentage of black voting-age population (VAP) over the entire periods. To operationalize the racial regime, I employ cutoff points of 40% and 65%, respectively, the latter of which is based on empirical and legal arguments about majority minority districts (Cameron, Epstein and O'halloran, 1996). While the first cutoff point is rather arbitrary, I performed robustness checks to confirm that my results do not depend on the marginal changes in the cutoff points. Among 2037 elections, 63.5% are from majority white districts, 25.6% are from racially balanced districts, and 11.4% are from majority minority districts, respectively. I also demonstrate that almost no district experienced drastic demographic changes and thus using the average percentage of black VAP seems valid (see Online Appendix B.2).

[15] When no candidate obtains the majority votes, then, the top two candidates compete each other in a run-off election. I found that 10% of the initial data were such run-off elections. I drop these elections to enable a better test for the model predictions since studying minority candidate emergence in run-offs faces with the same problem we have in general elections.

The outcome variable of interest is a binary variable denoting whether an election features any black candidate. At least one black candidate appears in about 26% of elections in the entire state, 6% of elections in majority white districts, 45.7% of races in racially balanced districts, and 94.2% of contests in majority minority districts. Figure 2 displays the distribution of the outcome variable across districts over time sorted by the average percentage of black voting-age population.

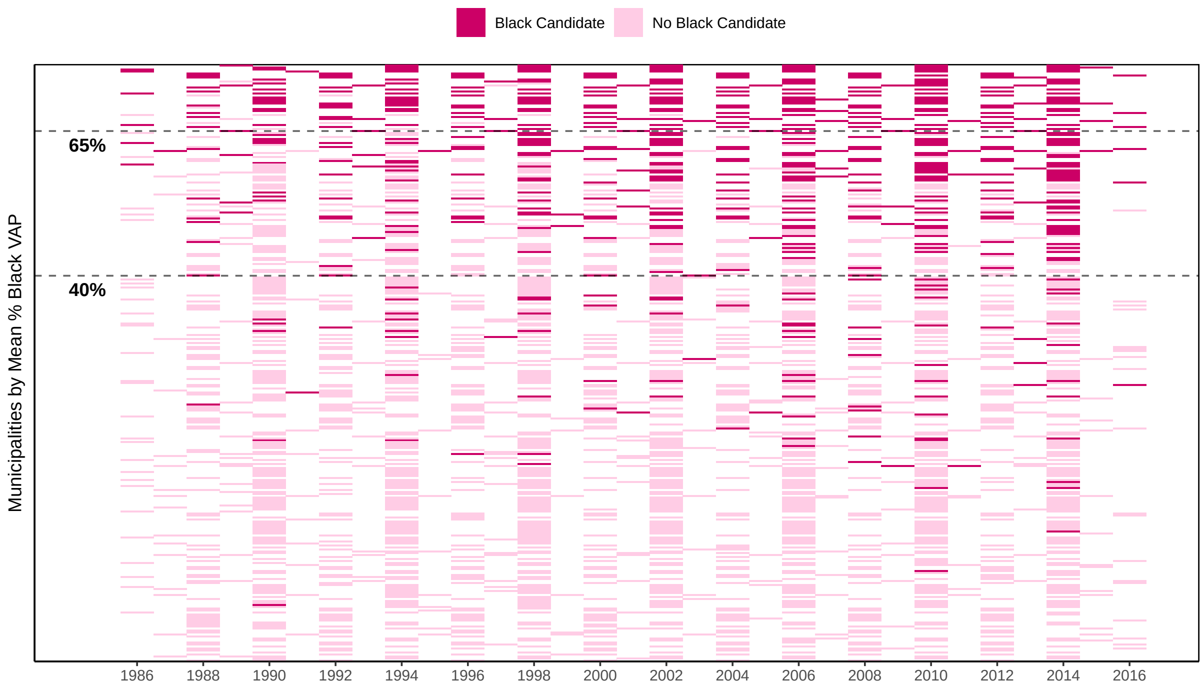

Figure 2: **Distribution of the Outcome Variable across Districts over Time**
Note: This figure portrays the distribution of the outcome variable in 303 districts from 1986 to 2016. The dark areas represent elections with one or more black candidates and the light areas indicate elections without any black candidate. The districts are ordered by the average percentage of black voting-age population in the entire time period from the highest (top) to the lowest (bottom).

As discussed in the "theory" section, the racial margin of victory is measured as $M = \frac{1}{2}(V_{t-1}^{M} - V_{t-1}^{W}) + 50$, where V_{t-1}^{M} and V_{t-1}^{M} are the vote shares obtained by the top minority and white candidates, respectively. Substantively, it represents the level of vote shares that need to be modified in order to change the race of the winner. Online Appendix B provides a summary of descriptive statistics of the data.

5 Validation of the Logical Model

5.1 Model Prediction and its Novelty

I begin by examining the predictive power of the logical model. Specifically, I compute the predicted values for black candidate emergence based only on Equation (4) and compare them to the observed data points. Figure 3 visualizes the model predictions along with observed data as a contour plot. The contour plot takes the racial margin of victory on its x-axis and the percentage of black VAP on its y-axis.

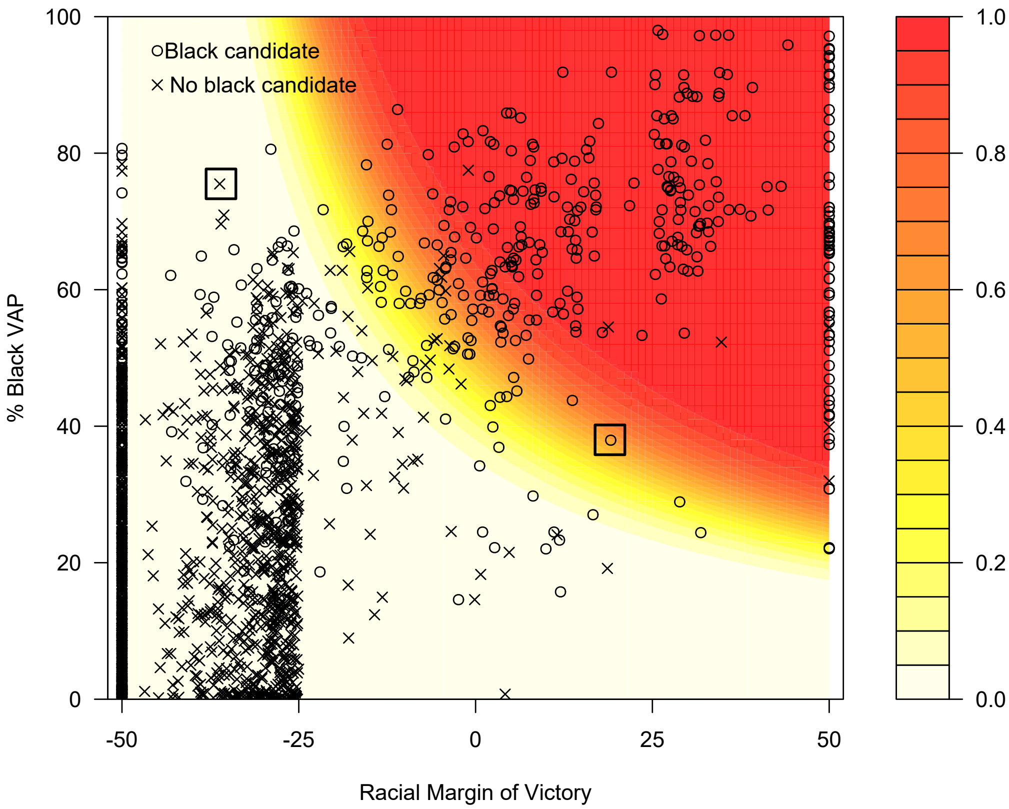

Figure 3: **Model Prediction of Black Candidate Emergence**
Note: This figure visualizes the predicted probabilities of black candidate emergence as a function of racial electoral performance and % black VAP in a contour plot. Elections with black candidates are shown as open circles (○), whereas elections without black candidate are represented by crosses (×). Two observations marked by squares are the cases that previous research cannot explain but the logical model can.

The gradation of color represents the predicted probability of black candidate emergence based on the logical model. Thus, the redder areas show higher probabilities that black candidates running for office and whiter regions indicate lower probabilities. This plot implies that even a casual visual inspection can tell that the logical model predicts observed black candidate emergence quite accurately. Namely, most elections with black candidates (open circles) are located in the regions with predicted probabilities of 0.5 and higher and most elections without black candidate (crosses) appear in the areas with predicted probabilities of 0.5 and lower.

To highlight the contribution of the logical model, I mark (by squaring) two observations that previous research cannot fully explain, but the logical model can properly account for. The left square shows the absence of minority candidate in a majority minority district, whereas the right square marks the presence of minority candidate in a majority white district. Both cases are puzzling to conventional studies because it is usually believed that minority candidate emergence is heavily determined by the size of minority voters. Thus, both demand and supply-side theories suggest that we must observe minority candidate(s) running in the first case, and we should not expect any minority candidacy in the second case.

In contrast, the logical model indicates that both two cases "make sense." Indeed, the logical model expects the absence of minority candidates *even in* majority minority districts when their co-ethnic candidates performed "poorly" in the last elections (relative to their white counterparts), whereas it anticipates the emergence of minority candidates *even in* majority white districts when their co-ethnic candidates won the last elections with large racial margin of victories (performed "well" relative to the white candidates). The model predictions for both cases are 0.00 and 0.56, respectively. Thus, not only the logical model can explain such previously puzzling cases, but also it can provide a highly accurate prediction on the outcomes. In short, the logical model offers a novel insight to the literature that the district racial composition is only half the story of minority candidate emergence.

5.2 Predictive Performance

To further quantify the above finding, I compute the accuracy of model predictions based on the expected Percentage Correctly Predicted (ePCP). The ePCP provides the percentage of observations for which the model can correctly predict their values, while accounting for "how close" such predictions are (Herron, 1999). The results are reported in Table 1.

Column 1 shows that *only based on the logical model* we can correctly predict about 88% of cases in our data. Indeed, this (in-sample) prediction is quite accurate and it further provides supportive evidence

	In-sample Prediction (%)			Out-of-sample Prediction (%)		
	Logical Model	LPM	Logit	Logical Model	LPM	Logit
All Districts (N=2037)	**88.6**	83.2	85.0	**88.3**	82.9	84.7
Majority White Districts (N=1293)	**94.1**	90.3	90.8	**94.1**	89.7	90.1
Racially Balanced Districts (N=521)	**73.7**	66.4	67.0	**75.6**	66.9	67.7
Majority Minority Districts (N=223)	**90.9**	91.8	93.6	**91.1**	90.1	91.4

Table 1: **Predictive Performance of the Logical Model**
Note: This table reports the weighted percentage of observations for which the model correctly predicts their values (ePCP), where the weight reflects the distance between the true value and predicted value. ePCP is calculated as $\frac{1}{N}(\sum_{y_i=1}\widehat{P}_i + \sum_{y_i=0}(1-\widehat{P}_i))$, where y_i and $\widehat{P}_i$ are true and predicted values for unit i, respectively. ePCP was then multiplied by 100 to be in the percentage scale. The variance parameter is set to 1 in the logical model. For out-of-sample prediction, leave-P-Out cross validation is employed where P is the number of units from the same municipality.

for the logical model. It also reports that the logical model successfully predicted about 95%, 72%, and 87% of cases in majority white, racially balanced, and majority minority districts. These results are rather remarkable because, again, our prediction is based only on our logical model and not machine learning or other predictive models.

To further buttress this point, I run a series of linear probability models (LPM) with the proportion of blacks and racial electoral performance as predictors and generate ePCPs. The results are reported in Column 2 of Table 1. The ePCP based on this "atheoretical" regression is about 83% for the full sample and this is a lot lower than the result based on the theoretical model.[16] Moreover, the LPM generated lower ePCPs compared to the logical model for majority white and racially balanced districts. For majority minority districts, the LPM seems to perform slightly better than the logical model. However, since these districts consist of only about 11% of the data, the logical model performs better on average as shown in Row 1. Is this an artifact of the fact that I use the LPM? To investigate this concern, I also run a set of logistic regression models. The results are reported in Column 3. While logistic regressions seem to have higher ePCPs than LPMs, the overall patterns remain the same and they do not perform better than the logical model. This grants additional confidence that the theoretical model has higher predictive power than more complex statistical models.

Moreover, I also run LMP and logistic regressions with 34 additional variables that previous research tends to include in its statistical models (The detailed results are shown in Online Appendix C.3). I find that ePCPs based on theses extended regressions are still lower than the ePCPs from the logical model

[16] I call this model "atheoretical" because I did not provide any consideration on the functional form of the predictors.

despite the fact that the extended regressions have 34 additional variables and, importantly, four of them are statistically significant at the benchmark of 0.05. Theses results demonstrates the power of logical models and that parsimonious theoretical models can often better describe the world than *ad hoc* statistical models (Taagepera, 2008; Shugart and Taagepera, 2017).

Some readers may be concerned with the external validity of these findings. That is, they might wonder to what extent the results are stemmed from the property of the particular data that I am using. To resolve such concern, I compare the out-of-sample predictive performance of the logical model and LPM and logistic regressions based on Leave-P-Out cross validation (Celisse et al., 2014). The idea is that we can simulate what the model's predictive performance looks like outside this particular data by evaluating the predictive performance on P subsets of the original data, where I use the number of municipalities for P. Columns 4-6 of Table 1 illustrate that the out-of-sample performance of the logical model is still quite high (88.3%) and higher or equal to the two regressions, suggesting that the logical model does not just fit this data by chance but provides a stable prediction outside it.

5.3 Internal Validity

If the logical model indeed predicts minority candidate "emergence" (E) through the likelihood of winning ($\widehat{P}$), it *should not* predict other political outcomes such as the "number" of minority candidates running for office, whereas it *must* predict minority candidate victory. This is a critical point for validating the logical model since my argument is based on the strategic entry of the top minority candidates. Although it is extremely difficult to verify this aspect of the theory, it is still possible to examine some observable implications regarding this story. Such observable implications include (1) the majority of elections with minority candidate emergence feature only one minority candidate; (2) the logical model does not predict the number of minority candidates; (3) the number of minority candidates follows some random pattern; and (4) the logical model predicts minority candidate victory. I found evidence for all of the four expectations.

First, I found that 60% of the elections where black candidates appear had only one black candidate, followed by 26% with two black candidates, 10% with three black candidates, and 4% with four or more black candidates. Second, I found that there is not clear pattern between the model predictions and the distribution of the number of black candidates. The left panel of Figure 4 plots the number of black candidates on ballots over the model predictions (as contour plot). Just a visual inspection can tell that there is no clear relationship between the number of black candidates and the racial margin of victory (x-axis). While the number of black candidates seem to be associated with the percentage of black VAP (y-axis), this may be mostly attributed to the fact that two or more black candidates rarely appear in elections with 50% or lower black VAP. To submit more quantitative evidence, I also run a series of count models with the number of black candidates as the dependent variable and the racial margin of victory and the percentage of black VAP as two explanatory variables (The results are reported in Online Appendix C.1). I found that even though the percentage of black VAP is positively associated with the number of black candidates, the racial margin of victory has no association that is statistically (and substantively) significant.

Third, I found that the number of black candidates follows a power law (in its tail) (the right panel of Figure 4).[17] The power law can be observed in diverse natural and social phenomena where many observations are clustered around some typical values (Clauset, Shalizi and Newman, 2009). In other words, when empirical data follows the power law (distribution), very few values are observed with significantly high frequencies with many other values with low frequencies (e.g., the intensity of wars, the severity of terrorist attacks, the frequency of US family names, the number of human population in US cities, etc). The right plot of Figure 4 visualizes this pattern for the number of black candidates in our data. The hypothesis test suggests that I cannot reject the null hypothesis that the data follows a power law distribution (p-value=0.79).

Finally, I found that the logical model can correctly predict 94% of black candidate victory (in ePCP). This result is rather stunning. From a theoretical perspective, this means that minority candidates indeed decide whether to run for office by carefully consulting with the likelihood of winning (and the cost and befit terms have less importance in the candidate entry model as discussed in Online Appendix A). From a logical model standpoint, the result is critical since it enables us to approximate the logical model of minority descriptive representation by the logical model of minority candidate emergence. Such approximation contributes to the ultimate goal of logical models where few institutional variables that are subject to change by electoral engineering (e.g., the size of minority voters via redistricting and margin of victory via electoral rule) can predict the outcomes of interest (e.g., the number of minority representatives) (Taagepera, 2007; Li and Shugart, 2016; Shugart and Taagepera, 2017).

These evidence implies that the logical model can withstand tests for its theoretical consistency.[18]

[17] In reality, we can never be certain that the observed quantities follow a power law due to the sampling variability in tail values. Thus, a more conservative conclusion is that, given the performed hypothesis test, it is more likely that the tail of the number of black candidates is drawn from a power-law distribution (with the minimum number is 3).

[18] I further conduct other placebo tests with female candidate emergence as false outcomes for the model prediction (Online Appendix C.2), finding that the logical model does not predict the false outcomes. This suggests that the logical model is indeed a model of "minority" candidate emergence.

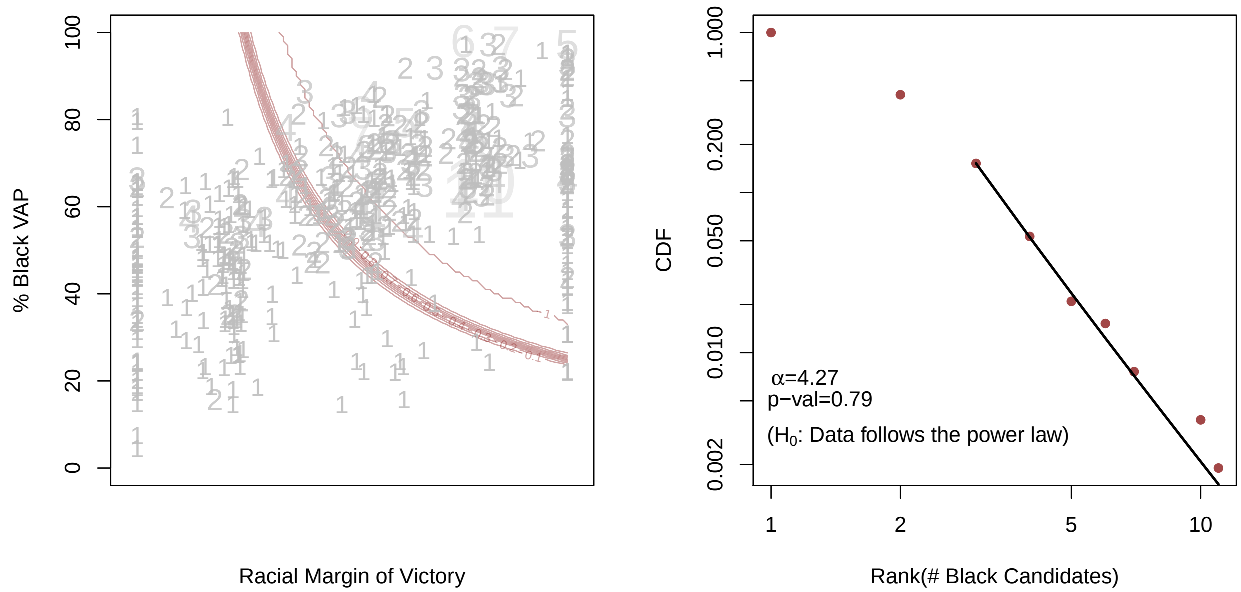

Figure 4: **The Number of Black Candidates Does Not Follow the Model Prediction, but a Power Law**
Note: The left panel plots the number of black candidates running for office over the parameter space of the logical model (with model predictions as contour plots). The right panel shows that the number of black candidates follows a power law distribution.

5.4 Regression Analysis

The above analysis demonstrates that, *all things being equal*, the logical model can predict black candidate emergence as a function of the two variables with higher predictive performance. It is important to acknowledge, however, that such conditional statement with the logical model is somewhat different from conditional statements (i.e., "ceteris paribus") used in more conventional regression analysis and causal inference, respectively (Clark and Golder, 2015). Indeed, some readers may wonder if the observable implications from the logical model hold even after controlling for other variables as high predictive performance does not necessarily mean the presence of causal mechanism (Breiman et al., 2001; Shmueli et al., 2010).

Here, I tackle this problem by defining the racial margin of victory as the treatment variable and black candidate emergence as the outcome variable of interest. I choose to focus on the racial margin of victory because it is a novel concept in the literature, whereas it is widely known that the size of minority voters is strongly correlated with minority candidate emergence and victory (e.g., Cameron, Epstein and O'halloran, 1996; Shah, 2014; Juenke, 2014) and thus re-discovering such correlation is not surprising. Thus, I specifically test the two hypotheses derived in Section 3 and visualized in the right plot of Figure 1: the increase in the racial margin of victory raises the probability of minority candidate emergence (H_1) and the impact of the racial margin of victory becomes the greatest when the district is more racially balanced (H_2). Due to the nature of our problem with the continuous treatment, binary outcome, large number of municipalities (N=301), large number of time periods (T≈30), and the simultaneous nature of the logical model (i.e., multiplicative term of $(MC)^{1/2}$), I apply a more conventional regression analysis rather than the causal inference approach based on the potential-outcomes framework (Holland, 1986; Hernan and Robins, 2010; Morgan and Winship, 2015). Indeed, integrating the logical model approach with the potential-outcomes framework is one of the promising areas of inquiry (and beyond the scope of this paper) and I leave to future research the development of such framework.

I first consider the bivariate relationship between the racial margin of victory and minority candidate emergence using a Bayesian logistic regression. The results are shown in column 1 of Table 2. The results represent that a higher racial margin of victory is associated with a greater chance of observing minority candidates. To account for a set of potential confounders which both affect the racial margin of victory (at time $t-1$) and minority candidate emergence (at time t) as described in a directed acyclic graph in Online Appendix **??**, I next include several covariates including the proportion of black VAP, the proportion of black and white population with B.A., and election cycle and apply district and year specific random intercepts to account for other unobserved potential confounders which are specific to each district and year. The results shown in column 2 indicate that the relationship between the racial margin of victory and minority candidate emergence still holds, supporting Hypothesis 1.

To test Hypothesis 2, I next estimate a set of Bayesian logistic regression by featuring random coefficients for racial electoral performance by the three types of districts or the racial regime (i.e., majority white, racially balanced, and majority minority districts). Column 3 reports the results, demonstrating that in all racial regime racial electoral performance is positively associated with the likelihood of observing minority candidate. To facilitate the interpretation, I plot the predicted probabilities of minority candidate emergence against the racial margin of victory by the racial regime based on the results on Column 3 (in Figure 5). Here, I calculate the predicted probabilities using median values of covariates for each racial regime to make a more realistic comparison across racial regime.

	Column (1)	Column (2)	Column (3)	Column (4)
Racial margin	.085	.048		
	(.077, .093)	(.039, .058)		
Racial margin (majority white)			.039	.065
			(.025, .053)	(.042, .089)
Racial margin (racially balanced)			.051	.078
			(.039, .063)	(.059, .098)
Racial margin (majority black)			.062	.082
			(.039, .087)	(.040, .128)
Covariates		✓	✓	✓
Random effects (28 years)		✓	✓	✓
Random effects (301 units)		✓	✓	✓
N	2037	2037	2037	1169

Table 2: **Regression Results**
Note: This table reports the posterior estimates for the coefficients of interest in Bayesian logistic regression. 95% credible intervals are shown inside parentheses. Rows 2-4 report that random slopes for racial margin of victory are estimated by the types of districts.

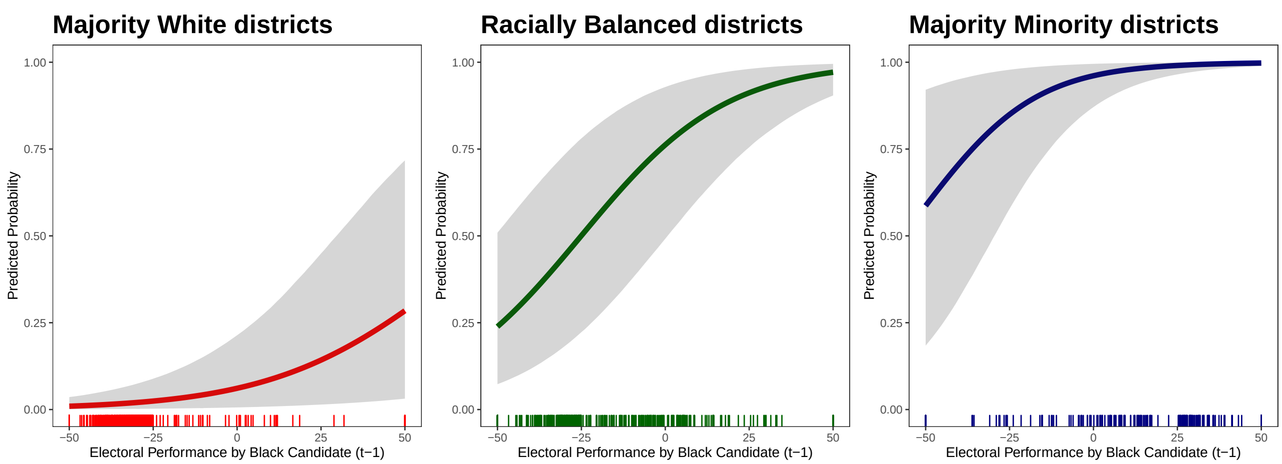

Figure 5: **The Substantive Effects of the Racial Margin of Victory on Minority Candidate Emergence**
Note: This figure displays predicted probabilities (bold curves) that black candidates running for office under three racial regimes with 95% credible intervals (gray shades). Rug plots are added to show the empirical distributions of the racial electoral performance variable in my data. Predicted probabilities are computed based on typical covariate values for each type of district (i.e., each type has a different baseline).

If the logical model has a valid explanatory power, I expect to see a similar picture to the simulated results in the right panel of Figure 1. The predicted probabilities presented in Figure 5 show that this is exactly the case. The figure illustrates that the substantive effect of the racial margin of victory (i.e., the difference in predicted probabilities evaluated at the minimum and maximum racial electoral performance) is larger in racially balanced districts than in majority white or majority black districts. In racially balanced districts, the change in the racial margin of victory from the lowest to the highest seems to boost the predicted probability about 0.8, whereas the change in the same condition seems to be more modest and about 0.2 in majority white districts and 0.4 in majority black districts. This provides a strong evidence for Hypothesis 2 and the theoretical story behind it that the value of the racial margin of victory (at time $t-1$) as a heuristic becomes the greatest when the racial composition is not informative (C=50). I performed various types of robustness checks to examine the interval validity of the above analysis (Online Appendix C.4), finding that the substantive conclusion stays the same.

6 Concluding Remarks

Minority descriptive representation is a critical component of representative democracy and has been extensively studied by scholars of racial politics, representation, voting rights, election law, and electoral systems. This article contributes to the rich literature by providing a quantitatively predictive logical model of minority candidate emergence (and by proxy minority descriptive representation), which unifies previously divided theoretical camps of voter demand and candidate supply; models minority candidate decision-making based on deductive logic and clearly stated assumptions; and offers a highly accurate prediction of minority candidate emergence and victory. The logical model introduced a novel concept of the racial margin of victory and demonstrated that the district racial composition is only half of the story, enabling us to explain the outcomes that previous research was not able to account for (e.g., minority representatives from majority white districts and no minority candidates from majority minority districts).

The logical model derived in this article offers a useful benchmark for developing logical models of minority representation more broadly. If researchers can logically connect institutional variables that are subject to change with the relevant outcomes of minority representation (e.g., the number of minority representatives in a legislature) and represent the relationship as a simply mathematical equation, it becomes possible for them to design a polity in such a way that promotes minority representation, enabling "applied" political science (Li and Shugart, 2016). Developing and linking such logical models makes a fruitful opportunity for future research.

As for any novel model, the proposed model has limitations. Thus, future research must further validate the model prediction with diverse data sets, different minority groups of interest, and alternative levels of elections. Moreover, a potentially promising topic for future research would include (but not be limited to) the derivation of the racial margin of victory in electoral systems other than first-past-the-post elections such as the single-winner and multi-winner ranked-choice voting.

References

Abosch, Yishaiya, Matt A Barreto and Nathan D Woods. 2007. “An assessment of racially polarized voting for and against Latino candidates in California.” *Voting Right Act Reauthorization of 2006: Perspectives on Democracy, Participation, and Power* pp. 107–131.

Aldrich, John H. 1995. *Why parties?: The origin and transformation of political parties in America*. University of Chicago Press.

Barreto, Matt A, Gary M Segura and Nathan D Woods. 2004. “The mobilizing effect of majority–minority districts on Latino turnout.” *American Political Science Review* 98(1):65–75.

Bedoya, Alvaro. 2005. “The Unforseen Effects of Georgia v. Ashcroft on the Latino Community.” *Yale LJ* 115:2112.

Black, Gordon S. 1972. “A theory of political ambition: Career choices and the role of structural incentives.” *American political science review* 66(1):144–159.

Breiman, Leo et al. 2001. “Statistical modeling: The two cultures (with comments and a rejoinder by the author).” *Statistical science* 16(3):199–231.

Bullock III, Charles S and A Brock Smith. 1990. “Black success in local runoff elections.” *The Journal of Politics* 52(4):1205–1220.

Bullock III, Charles S and Loch K Johnson. 1985. “Runoff elections in Georgia.” *The Journal of Politics* 47(3):937–946.

Cameron, Charles, David Epstein and Sharyn O’halloran. 1996. “Do majority-minority districts maximize substantive black representation in Congress?” *American Political Science Review* 90(4):794–812.

Canon, David T. 1999. *Race, redistricting, and representation: The unintended consequences of black majority districts*. University of Chicago Press.

Casellas, Jason P. 2010. *Latino representation in state houses and congress*. Cambridge University Press.

Celisse, Alain et al. 2014. “Optimal cross-validation in density estimation with the $L\hat{}\{2\}$-loss.” *The Annals of Statistics* 42(5):1879–1910.

Clark, William Roberts and Matt Golder. 2015. “Big data, causal inference, and formal theory: Contradictory trends in political science?: Introduction.” *PS: Political Science & Politics* 48(1):65–70.

Clauset, Aaron, Cosma Rohilla Shalizi and Mark EJ Newman. 2009. “Power-law distributions in empirical data.” *SIAM review* 51(4):661–703.

Colomer, Josep M. 2007. “What other sciences look like.” *European Political Science* 6(2):134–142.

Cox, Gary W. 1997. *Making votes count: strategic coordination in the world’s electoral systems*. Cambridge University Press.

Elmendorf, Christopher S, Kevin M Quinn and Marisa A Abrajano. 2016. “Racially Polarized Voting.” *U. Chi. L. Rev.* 83:587.

Fox, Richard L and Jennifer L Lawless. 2004. "Entering the arena? Gender and the decision to run for office." *American Journal of Political Science* 48(2):264–280.

Fraga, Bernard, eric Gonzalez Juenke and Paru Shah. 2019. "One Run Leads to Another: Minority Incumbents and the Emergence of Lower Ticket Minority Candidates.".

Fraga, Bernard L. 2014. Race, Party, and Candidate Prospects across the Multiple Stages of Congressional Elections. Technical report Working Paper. http://www. bernardfraga. com/uploads/2/2/3/4/22341374 .

Granato, Jim, Melody Lo and MC Sunny Wong. 2010. "A framework for unifying formal and empirical analysis." *American Journal of Political Science* 54(3):783–797.

Grofman, Bernard, Lisa Handley and Richard G Niemi. 1992. *Minority representation and the quest for voting equality*. Cambridge University Press.

Hernan, Miguel A and James M Robins. 2010. "Causal inference.".

Herron, Michael C. 1999. "Postestimation uncertainty in limited dependent variable models." *Political Analysis* 8(1):83–98.

Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American statistical Association* 81(396):945–960.

Jacobson, Gary C and Samuel Kernell. 1983. "Strategy and choice in congressional elections.".

John, Sarah, Haley Smith and Elizabeth Zack. 2018. "The alternative vote: Do changes in single-member voting systems affect descriptive representation of women and minorities?" *Electoral Studies* 54:90–102.

Jones, Bryan D. 1999. "Bounded rationality." *Annual review of political science* 2(1):297–321.

Juenke, Eric Gonzalez. 2014. "Ignorance is bias: The effect of Latino Losers on models of Latino representation." *American Journal of Political Science* 58(3):593–603.

Juenke, Eric Gonzalez and Paru Shah. 2015. "Not the usual story: the effect of candidate supply on models of Latino descriptive representation." *Politics, Groups, and Identities* 3(3):438–453.

Juenke, Eric Gonzalez and Paru Shah. 2016. "Demand and supply: Racial and ethnic minority candidates in White districts." *Journal of Race, Ethnicity and Politics* 1(1):60–90.

Keele, Luke J, Paru R Shah, Ismail White and Kristine Kay. 2017. "Black Candidates and Black Turnout: A Study of Viability in Louisiana Mayoral Elections." *The Journal of Politics* 79(3):780–791.

King, Gary, Robert O Keohane and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research*. Princeton university press.

Lazarus, Jeffrey. 2008. "Buying in: Testing the rational model of candidate entry." *The Journal of Politics* 70(3):837–850.

Li, Yuhui and Matthew S Shugart. 2016. "The seat product model of the effective number of parties: A case for applied political science." *Electoral Studies* 41:23–34.

Lublin, David. 1999. *The paradox of representation: Racial gerrymandering and minority interests in Congress*. Princeton University Press.

Lublin, David, Thomas L Brunell, Bernard Grofman and Lisa Handley. 2009. "Has the Voting Rights Act Outlived Its Usefulness? In a Word,No." *Legislative Studies Quarterly* 34(4):525–553.

Magrino, Thomas R, Ronald L Rivest, Emily Shen and David Wagner. 2011. Computing the Margin of Victory in IRV Elections. In *EVT/WOTE*.

Maisel, L Sandy and Walter J Stone. 1997. "Determinants of candidate emergence in US house elections: An exploratory study." *Legislative Studies Quarterly* pp. 79–96.

Marschall, Melissa J, Anirudh VS Ruhil and Paru R Shah. 2010. "The new racial calculus: Electoral institutions and black representation in local legislatures." *American Journal of Political Science* 54(1):107–124.

Marschall, Melissa and Paru Shah. 2013. "The Local Elections in America Project [database].".
URL: *http://www.leap-elections.org/*

Morgan, Stephen L and Christopher Winship. 2015. *Counterfactuals and causal inference*. Cambridge University Press.

Reed, Steven R. 1990. "Structure and behaviour: Extending Duverger's Law to the Japanese case." *British Journal of Political Science* 20(3):335–356.

Shah, Paru. 2014. "It takes a Black candidate: A supply-side theory of minority representation." *Political Research Quarterly* 67(2):266–279.

Shah, Paru. 2017. "Racial change, racial threat, and minority representation in cities." *Urban Affairs Review* p. 1078087417737386.

Shah, Paru R and Nicholas R Davis. 2017. "Comparing Three Methods of Measuring Race/Ethnicity." *Journal of Race, Ethnicity and Politics* 2(1):124–139.

Shmueli, Galit et al. 2010. "To explain or to predict?" *Statistical science* 25(3):289–310.

Shugart, Matthew S and Rein Taagepera. 2017. *Votes from seats: Logical models of electoral systems*. Cambridge University Press.

Stone, Walter J and L Sandy Maisel. 2003. "The not-so-simple calculus of winning: Potential US house candidates' nomination and general election prospects." *The Journal of Politics* 65(4):951–977.

Stone, Walter J, L Sandy Maisel and Cherie D Maestas. 2004. "Quality counts: Extending the strategic politician model of incumbent deterrence." *American Journal of Political Science* 48(3):479–495.

Taagepera, Rein. 2007. *Predicting party sizes: The logic of simple electoral systems*. Oxford University Press on Demand.

Taagepera, Rein. 2008. *Making social sciences more scientific: The need for predictive models*. OUP Oxford.

Taagepera, Rein and Miroslav Nemčok. 2019. "Majoritarian and consensual democracies: The seat product connection." *Party Politics* p. 1354068819871441.

Trounstine, Jessica and Melody E Valdini. 2008. "The context matters: The effects of single-member versus at-large districts on city council diversity." *American Journal of Political Science* 52(3):554–569.

Xia, Lirong. 2012. Computing the margin of victory for various voting rules. In *Proceedings of the 13th ACM Conference on Electronic Commerce*. ACM pp. 982–999.

Online Appendix

For "Unifying the Demand and Supply-Side Theories of Minority Descriptive Representation: A Logical Model"

Table of Contents

A Derivation of the Logical Model **2**

A.1 Candidate Entry . . . 2

A.2 Racial Margin of Victory . . . 3

A.3 Two Heuristics for the Racial Margin of Victory . . . 4

A.4 Proof on Racial Margin of Victory . . . 6

A.5 Additional Discussion on the Logical Model . . . 9

B Louisiana Mayoral Elections **11**

B.1 Distributions of M and C . . . 11

B.2 Stability of the Racial Regime . . . 11

B.3 Overtime Variation of Black Candidate Emergence . . . 11

B.4 Overtime Variation of the Racial Margin of Victory . . . 12

B.5 Registration by Race . . . 13

C Additional Findings **15**

C.1 Count Model Regression Results . . . 15

C.2 Model Prediction with Placebo Outcomes . . . 15

C.3 Extended Regression Results . . . 15

C.4 Robustness Checks for Bayesian Logistic Regressions . . . 15

A Derivation of the Logical Model

In this section, I detail the derivation of the logical model of minority candidate emergence $\text{Pr}(E) = \widehat{P} = \mathbf{F}\Big((MC)^{1/2} - 50\Big)$. This section also accompanies with the proof on the racial margin of victory and additional discussion on the logical model.

A.1 Candidate Entry

Based on Assumptions 1-4, I consider the logical model of minority candidate emergence. Following the literature of political ambition (Black, 1972; Lazarus, 2008; Aldrich, 1995; Jacobson and Kernell, 1983; Stone, Maisel and Maestas, 2004), I start by applying the following model of candidate entry to minority candidate emergence:

$$u_{it} = \widehat{P}_{it} B_{it} - C_{it} \tag{1}$$

Here, u_{it} is the utility that candidate i obtains from running for office, $\widehat{P}_{it}$ is the candidate's estimate of the probability of winning if they enter the race, B_{it} is the benefits of holding the office, and C_{it} is the cost of running, which is greater than zero, all at time t.[1] Black (1972) suggests that a rational office-seeking candidate enters the race if and only if $\widehat{P}_{it} B_{it} > C_{it} > 0$ and $u_{it} > u_{it}(A)$, where $u_{it}(A)$ expresses the benefit that the candidate receives when staying out.[2]

While this model looks fairly simple at the first glance, it is in fact a very complicated model where four variables and three inequalities collectively determine the outcome. As we often prefer parsimonious models to complex models in logical model building (Taagepera, 2008), here, I only focus on the win probability term by taking other terms out of the equation.[3] Consequently, I directly model the probability of entry E_{it} as the win probability term as:

$$\text{Pr}(E_{it}) = \widehat{P}_{it} \tag{2}$$

Substantively, this implies that minority candidate emergence can be predicted by knowing about $\widehat{P}_{it}$, that is, how likely it is for minority candidates to win when they enter the race. The primary argument that I am making here is that *all things being equal* minority candidate emergence can be solely predicted by such win probability. The second half of model building then involves sources of such win probability.

[1] In contrast to previous studies, I use the notation $\widehat{P}_{it}$ (as opposed to P_{it}) to make it clear that it is a candidate's estimate of the true winning probability.

[2] Lazarus (2008) discusses that $u_{it}(A)$ is usually very different for experienced politicians and amateurs, which affect the required level of $\widehat{P}_{it}$ in the model, but I do not distinguish the two types of politicians here.

[3] The resulting model is also obtained by assuming that $B_{it} = 1, C_{it} = 0.5$, and $u_{it}(A) \sim N(0, \sigma^2)$, where $\sigma^2 > 0$. Under these assumptions, the original model tells that minority candidates run if and only if $\widehat{P}_{it} > 0.5 + u_{it}(A)$. Since 0.5 and u_{it} are constant and random numbers, we can say that the probability of entry is a function of the win probability term as $\text{Pr}(E_{it}) = \widehat{P}_{it} + u'_{it}$, where u'_{it} combines the constant and the error term. If we ignore the random component, we obtain Equation (2). While the model can be presented both in a deterministic form (without the error term) and stochastic form (with the error term), I choose to use the deterministic form due to simplicity.

A.2 Racial Margin of Victory

How can minority candidates calculate $\widehat{P}_{it}$? Here, I argue that the win probability is a function of what I call the *racial margin of victory*. In order to elaborate what it means, let me first introduce the concept of the general margin of victory as follows:

DEFINITION 1 (General Margin of Victory): *The general margin of victory is the minimum number of vote shares that have to be modified in order to change the outcome (i.e., winner) of an election.*[4]

In first-past-the-post (FPTP) elections, the margin of victory is computed as half the difference in vote shares between the winner and the runner-up. For example, when I have a winner who received 60% of ballots and a runner-up who obtained 40% of ballots, the general margin of victory is $\frac{1}{2} * (60 - 40) = 10$. Thus, if I remove 10% of ballots from the winner and allocate them to the runner-up, the election will be a tie (50% ballots vs. 50% ballots) and thus change the outcome of the election.[5]

Building on this concept, I then introduce the "racial" margin of victory as follows:

DEFINITION 2 (Racial Margin of Victory): *The racial margin of victory is the minimum number of vote shares that have to be modified in order to change the race of the winner of an election.*

To contextualize my argument, let me define African American voters as the minority group of interest and white voters as the majority group of reference. In first-past-the-post elections, then, the racial margin of victory (assuming no ties and rounding the difference) can be computed as below.

$$\text{Racial Margin of Victory} = \left|\frac{V^B - V^W}{2}\right|,$$

where V^B represents the number of **v**otes received by the top **b**lack candidate, V^W represents the number of **v**otes received by the top **w**hite candidate, and $|\cdot|$ is an absolute value operator. In other words, the racial margin of victory is calculated as half the absolute difference in the number of ballots received by the top black candidates and the number of ballots received by the top white candidates.[6]

The primary property of the racial margin of victory is that it does not distinguish the direction of racial change. Put differently, the concept is agnostic about whether the new winner is black or white after modifying the number of ballots. Removing the absolute value operator from it, however, enables us to calculate the minimum number of ballots that need to be modified in order to *replace a black winner with a white winner*. I call it as a *signed racial margin of victory* and define it as follows:

[4] The more accurate definition of the general margin of victory considers the minimum number of ballots instead of vote shares. **Add justification.**

[5] To compute the "tie-free" margin, I simply add 1 to the above number (Xia, 2012; Magrino et al., 2011), while I assume that there is always a tie breaker in my argument as discussed above.

[6] Here, I do not consider the presence of run-off elections and assume that every election is a decisive election. Surprisingly, this measurement only depends on the ballots received by the top black and white candidates and is not a function of other factors such as the number of black and white candidates, the internal distribution of ballots within the same racial group, and relative advantages of black runner-up to white runner-up. In Online Appendix A.4, I consider all eight possible scenarios (i.e., from one black candidate with one white candidate to multiple black candidates with no white candidate) and demonstrate that only vote shares of the top black and white candidates are required to calculate the racial margin of victory.

$$\text{Signed Racial Margin of Victory} = \frac{V^B - V^W}{2}$$

Thus, the signed racial margin of victory concisely represents the extent to which minority candidates safely secure their descriptive representation relative to their white counterparts. For simplicity, I hereafter use the margin of victory to refer to the signed racial margin of victory.

Now, how is the racial margin of victory tied back to the argument of minority candidate entry? The core idea is here that the probability of winning can be expressed as a deterministic function of the racial margin of victory. Thus, if black candidates are fully rational with complete information, the probability of winning becomes a function of the vote shares of the top black and white candidates in *upcoming* elections:

$$P_{it} = \mathbb{1}\left(\frac{V_{it}^B - V_{it}^W}{2} > 0\right) \tag{3}$$

Here, $\mathbb{1}$ is called an indicator function that takes 1 if the inside condition is satisfied and 0 otherwise. Substantively, this means that the probability of winning an electoral contest is 1 (i.e., winning with certainty) if the top black candidate obtains higher vote shares (or one more vote) than the top white candidate and 0 (i.e., losing with certainty) if the reverse is true.[7] Hence, if such future vote shares are *ex ante* known, the win probability is a deterministic quantity, and the rational minority candidate decides to run (not to run) for office with certainty.[8]

A.3 Two Heuristics for the Racial Margin of Victory

Although the above argument makes sense on the logical ground, it is usually impossible for any candidate to calculate future vote shares with no measurement error due to the cognitive burden, lack of enough information, and time restriction in decision making. Thus, I instead argue that the win probability is, in practice, calculated based on an educated guess about future vote shares.[9] I then assume the following win probability:

$$\widehat{P}_{it} = \mathbf{F}\left(\frac{\widehat{V}_{it}^B - \widehat{V}_{it}^W}{2}\right), \tag{4}$$

where the "hat symbols" ($\widehat{\cdot}$) imply that these vote shares are speculated quantities as opposed to known measurements. Here, **F** represents a cumulative distribution function (CDF) of any elliptical distribution (e.g., normal and t-distributions) with some mean and a finite variance and I use it for conceptual convenience (i.e., it is a monotonic function bounded by 0 and 1 and symmetric around the mean). One example of **F** is a CDF of a normal distribution with mean 0 and standard deviation 15, which maps the values between about -50 and 50 onto the probabilistic scale (i.e., numbers between 0 and 1). Here, the standard deviation controls the degree of non-zero probability that is assigned to values which are far away from the

[7]Here, I assume that there is always a tie-breaker and do not consider the case where the two quantities are the exactly same.

[8]Note that the argument presented so far is not surprising at all; rather, it simply clarifies a well known rational choice-type explanation that candidates decide to run for office when they see higher chances of winning. And if they can foreseen their electoral fortune, they will either enter the race or abstain. However, logical model building requires a clear definition and measurement of each concept and this section serves this purpose by connecting minority candidate emergence with a well discussed notion of the margin of victory in political science.

[9]This consideration is based on the idea of bounded rationality (Jones, 1999).

mean. *Substantively, it represents how much candidates would allow for the possibility of miscalculating the relative advantages of vote shares*.[10]

Next, I theorize that in order for black candidates, as for any candidates, to speculate the racial margin of victory they rely on **two different heuristics**, which include (1) the district racial composition (as an expected margin of victory) and (2) the racial margin of victory in the last elections. In other words, the probability of winning can be inferred from "what the district looks like now" and "what happened last time" in there. Let me explain what this means.

First, the district racial composition is informative because it can be directly translated to the expected difference in vote shares under the assumption of "perfect co-ethnic voting" (all voters vote for the candidate from their racial group). For example, if a black candidate is running against a white candidate in a district with 60% black and 40% white voters, the speculated vote share will become 60% and 40% for the black and white candidates, respectively; and the "best guess" for the difference in vote shares in the FTTP elections becomes 20% points (under the perfect co-ethnic voting assumption). Here, I call this quantity as an *expected* racial margin of victory (based on the district racial composition).

Second, the racial margin of victory in the last elections is instructive since it exactly tells how much black candidates can solicit crossover votes from white voters, while reserving co-ethnic votes from black voters, *if* the political climate remains the same from the last election. Put differently, the past racial margin of victory quantifies a potential deviation from the perfect co-ethnic voting assumption. It must be highlighted that the past racial margin of victory in the last election is a known (and thus measurable) quantity, while the racial margin of victory in an upcoming election is unknown (because nobody cannot see the future). To simplify the terminology, however, **I will simply call this quantity as the racial margin of victory (at time $t-1$) unless otherwise noted**.

Now, I formalize this argument of two heuristics by taking the "geometric mean" of the two quantities and replace the speculated racial margin of victory at time t with the geometric mean in Equation (4). Let $\tilde{C}$ denote the *expected* racial margin of victory (based on the district racial **C**omposition). Let $\tilde{M}$ denote the racial **M**argin of victory. Without changing their substantive meaning, I then shift both quantities by adding 50 to avoid the multiplication of negative numbers as follows:

$$M = \tilde{M} + 50 \tag{5}$$

$$C = \tilde{C} + 50 \tag{6}$$

Here, the "tilde symbols" ($\tilde{.}$) indicate that these are unadjusted (or raw) quantities. From now on, I call these adjusted quantities M and C simply as the racial margin of victory and the expected racial margin of victory, respectively unless otherwise noted.

Finally, connecting Equation (2) and Equation (4) along with the above argument yields:

$$\Pr(E) = \widehat{P} = \mathbf{F}\Big((MC)^{1/2} - 50\Big), \tag{7}$$

where $(MC)^{1/2}$ is the geometric mean of the two quantities and -50 is there to account for the adjustment shown in Equations (5)-(6). The purpose of this re-adjustment is to centralize the argument of **F** with mean

[10] My substantive argument does not change regardless of scaling as long as the function **F** has an appropriate tuning parameter as discussed below.

0 so that its scale does not change from Equation (4). Substantively, this ensures that positive (negative) numbers inside **F** mean higher (lower) probabilities of minority candidate emergence.

To supplement, $\tilde{M} = \frac{1}{2}(V_{t-1}^{B} - V_{t-1}^{W})$, where V_{t-1}^{B} and V_{t-1}^{M} are the vote shares obtained by the top black and white candidates, respectively at time $t-1$. $\tilde{C} = \%\text{black voters} - 50$ (at time t). Applying the adjustment yields:

$$M = \frac{1}{2}\Big(V_{t-1}^{B} - V_{t-1}^{W}\Big) + 50 \tag{8}$$

$$C = \%\text{black voters} \tag{9}$$

Thus, the adjusted expected racial margin of victory is algebraically equivalent to the percentage of black voters.

A.4 Proof on Racial Margin of Victory

In this subsection, I examine all possible different patterns of biracial elections and demonstrate that the racial margin of victory introduced in Online Appendix A.2 can be applied to any type of elections. Recall that the racial margin of victory refers to the minimum amount of vote shares one needs to modify in order to change the race of the winner in an election.

A.4.1 One black candidate, one white candidate

When there are only one black candidate and white candidate, respectively, on a ballot, all the possible patterns of electoral results can be obtained from the following permutations.

1st:	W	B
2nd:	B	W

In order to change the race of the winner, I am required to change the order of the winner and runner-up, which can be done at least with $|1/2 * (V_b - V_w)|$ votes, where V_b is the number of **v**otes received by **b**lack candidates and V_w the number of **v**otes obtained by **w**hite candidates. Since there are only one black candidate and only one white candidate (i.e., the black (white) candidate is always the top black (white) candidate), $|1/2 * (V_b - V_w)| = |1/2 * (V_{bt} - V_{wt})|$.

A.4.2 One black candidate, multiple whites candidates

When there are one black candidate and multiple white candidates, I only focus on the difference in the race of candidates. In other words, I do not distinguish one white candidate from another and only consider whether the ballot modification would lead to a winner with a different race from the original winner. Assuming that there are one black candidate and two white candidates, all the possible electoral outcomes are as follows.

1st: W W B

2nd: *W* B W
3rd: B *W* *W*

In the first pattern (column), what I focus on is the difference in the number of ballots received by the top white candidate and the black candidate. And this applies to all other two patterns. Thus, the racial margin of victory is computed as $|1/2 * (V_b - V_{wt})| = |1/2 * (V_{bt} - V_{wt})|$, since the black candidate on a ballot is always the top black candidate. Because the above quantity does not depend on the number of white candidates, the above equation remains the same even when I have more than three white candidates.

A.4.3 Multiple blacks candidates, one white candidate

In elections with two black candidates and only one white candidate, the possible ordering of candidates is shown as below.

1st: W B B
2nd: B W *B*
3rd: *B* *B* W

Following the same logic as the last case, the racial margin of victory is calculated as $|1/2*(V_{bt}-V_w)| = |1/2 * (V_{bt} - V_{wt})|$. Since this does not depend on the number of black candidates, it can be extended to elections with more than three black candidates.

A.4.4 Multiple black candidates, multiple white candidates

When there are two black and two white candidates on a ballot, the following covers all the possible electoral outcomes.

1st: W W W B B B
2nd: W B B W W B
3rd: B W B B W W
4th: B B W W B W

Here, it requires only $|1/2 * (V_{bt} - V_{wt})|$ to change the race of the winner. Since this is not a function of the number of black and white candidates, it can be extended to elections with more than three black or white candidates.

A.4.5 No black candidate, one white candidate

When an election is an unopposed election with a white winner, all the possible pattern is as follows.

1st: W

Now, this can be considered as a special case of 1. (One black candidate, one white candidate), where the black candidate received zero vote. Thus,

1st: W (V = 100)
2nd: B (V = 0)

Consequently, the racial margin of vote can be computed as $|1/2 * (V_b - V_w)| = |1/2 * (0 - V_{wt})| = 1/2 * V_{wt}$. And EPBC can be measured as $-1/2 * V_{wt}$.

A.4.6 No black candidate, multiple white candidates

When there are more than one white candidates with no black candidate, the possible electoral outcome can be demonstrated as follows:

1st: W
2nd: W
...

Now, this can be thought of as a special case of 2. (one black candidate, multiple white candidates) where the black candidate received zero vote. Thus,

1st: W
2nd: W
...
Last: B (V = 0)

Following the above explanation, the racial margin of victory can be computed as $|1/2 * (V_{bt} V_{wt})| = |1/2 * (0 V_{wt})| = 1/2 * V_{wt}$. Here, by definition, $1/2 * V_{wt}$ (as in 6.) $\leq 1/2 * V_{wt}$ (as in 5.).

A.4.7 One black candidate, no white candidate

In a case of unopposed election with a black winner, the possible electoral result is the mirrored version of 5. Thus,

1st: B

which can be rewritten as

1st: B (V = 100)
2nd: W (V = 0)

Therefore, the racial margin of victory can be computed as $|1/2 * (V_b - V_w)| = |1/2 * (V_{bt} - 0)| = 1/2 * V_{bt}$. And EPBC can be similarly measured as $1/2 * V_{bt}$.

A.4.8 Multiple black candidates, no white candidate

This is a mirrored version of 6. and all the possible patterns can be shown as:

1st: B
2nd: B
...
Last: W (V = 0)

Thus, the racial margin of victory can be computed as $|1/2*(V_{bt} - V_w)| = |1/2*(V_{bt} - 0)| = 1/2*V_{bt}$. By definition, $1/2 * V_{wt}$ (as in 8.) $\leq 1/2 * V_{wt}$ (as in 7.).

A.5 Additional Discussion on the Logical Model

In the above and the main text, I have demonstrated that minority candidate emergence can be modeled as the decision-making process of minority candidates who evaluate prior odds of winning in electoral contests. I have also provided supportive evidence for my argument by empirically testing observable implications from the model. In this subsection, I further discuss a broader implication of the model and resolve remaining theoretical issues.

First, some readers may wonder if the proposed model can be applicable to a broader set of electoral contests. I argue that the model is relevant for most elections conducted under the simple FPTP system with single-member districts. Such "simple" cases include many U.S. local elections, which are often non-partisan, and primary elections in various levels. In contrast, the proposed model cannot be directly applied to more complex systems, including the U.S. presidential primaries, general elections, and elections with different ballot structures such as ranked-choice voting and limited voting (John, Smith and Zack, 2018).

In these cases, the concept and calculation of margin of victory are often more complicated (Magrino et al., 2011; Xia, 2012) and it is relatively hard to theoretically explain the role of racial electoral performance as a useful heuristic for calculating the prior odds of winning. To understand minority candidate emergence in general elections is similarly challenging since the lack of minority candidates in general elections can mean either the absence of minority candidates *or* the defeat of minority candidates at the primary stage (Fraga, 2014). In the former case, minority candidate emergence at the general stage is a deterministic function of the one at the primary level, whereas it is not the case in the latter. Future research must expand the scope conditions to these more complex systems.

Next, the proposed model only accounts for minority candidate emergence in elections with two racial groups (Assumption 1). While the State of Louisiana fits the assumption, many districts in other states have more than two dominant racial groups. Thus, to expand the scope of the logical model to multiracial elections it is necessary to consider how to modify the concept of racial electoral performance in such environments. One potential approach is to think about biracial election as contests with one racial group and the other groups combined for each group. For instance, suppose that there are three groups of whites, hispanics and asians in a district. Then, racial electoral performance can be calculated for asian candidates based on past and expected racial margin of victories between the top asian candidate and top non-asian candidate. This strategy coincides with the current scholarship where researchers often consider the proportion of whites, blacks, hispanics, and asians as relevant predictors of minority candidate emergence (Fraga, Juenke and Shah, 2019). Future research must generalize the proposed model by incorporating multiple racial groups.

One potential confusion about the model could be that the model is taken as a model of *minority candidate success*. Therefore, it does not have any explanatory and predictive power if no minority candidate has ever run in a particular district. Moreover, the model cannot explain why the first minority candidate decided to run in a district. However, this is in fact an incorrect interpretation of the model. A more proper view is that the proposed model is a model of the "relative strength" of top minority candidate and top white candidate. To illustrate this point, let us consider an extreme case where there were ten white candidates and no black candidate at time $t-1$ in a district. Suppose that 80% and 20% of the electorate are white and black voters, respectively. Suppose also that the ten white candidates almost equally divided vote shares so that the vote share of the top white candidate was 10% plus few votes. Under this condition, racial electoral performance becomes approximately 5, which means that if a black candidate had been running in the district and obtained at least 5% plus few votes she could have been a winner. For this reason, the model states that minority candidates can emerge from majority white districts if electoral environments are favorable to them. As shown in Section 4 and Online Appendix A.2, the racial margin of victory can be calculated regardless of the presence of minority candidate at time $t-1$. Thus, the model can be applied to districts where no minority candidate has appeared before.

Finally, it may seem that the model does not account for the role of political parties or partisanship of voters. However, as emphasized in Section 3, I have referred to candidates *and* their teams as simply candidates. And I modeled the decision-making process of candidates and their teams altogether under the framework of bounded rationality. Thus, such decision-making can be also seen as the recruitment process by party stuff and leaders as some studies have considered (Shah, 2014; Juenke, 2014). On the other hand, it is true that the model does not explicitly include the effect of partisanship of candidates and voters on the decision-making process. However, such effect is indirectly incorporated in racial electoral performance at time $t-1$. If partisanship plays a role in minority candidate emergence such that parties attempt to push particular candidates to elect their own candidates, this strategic coordination must be reflected in racial electoral performance (Cox, 1997). That is, if there is any partisan coordination behind few candidates vote shares among candidates must have a skewed distribution, which in turn affects racial electoral performance. As a result, the role of parties is already included in the model under the stated assumptions. Nevertheless, it does not handle partisanship in the two-stage elections and such problem must be solved in future research.

B Louisiana Mayoral Elections

This section provides various descriptive statistics for Louisiana mayoral elections.

B.1 Distributions of M and C

In the Louisiana mayoral election data, the racial margin of victory is distributed somewhat irregularly, while the percentage of black voters seems to follow an exponential-type distribution. The Pearson's correlation coefficient between the two variables is 0.64, which can be visually confirmed as a strong positive correlation in the right panel below.

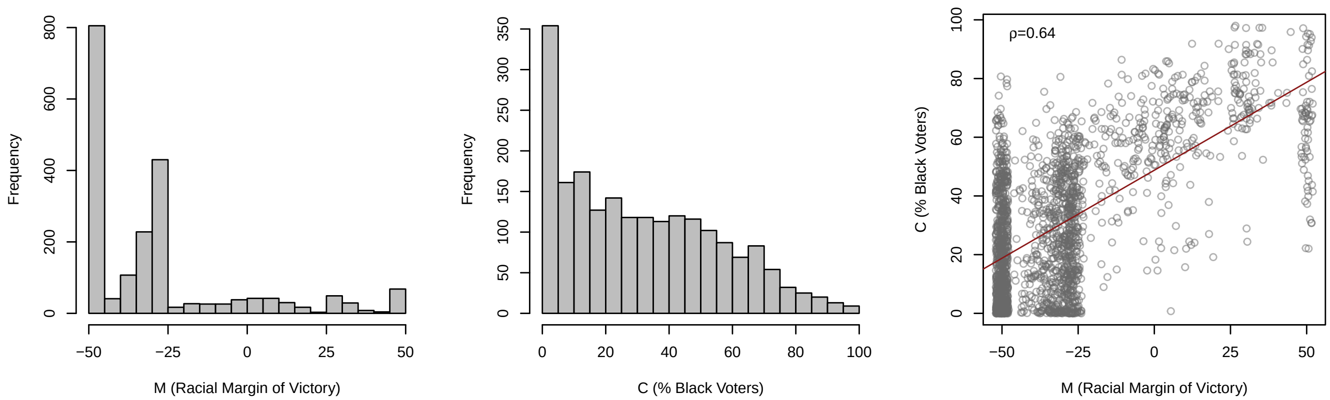

Figure B.1: **The Distributions of M and C with the Scatter Plot**
Note: The x-axis is jittered in the scatter plot.

B.2 Stability of the Racial Regime

To code the racial regime, I use the cutoff points of 40% and 65% (for the average black VAP). Figure B.2 shows the stability of such coding scheme. It portrays that very few municipalities "cross" the cutoff points within the observed periods.

B.3 Overtime Variation of Black Candidate Emergence

To investigate whether there is any temporal pattern in black candidate emergence, Figure B.3 plots the proportion of elections with black candidates over time. It also accounts for the racial regime. I find that there is almost no temporal pattern except for the fact that black candidate emergence has been slightly more common in racially balanced districts over time.

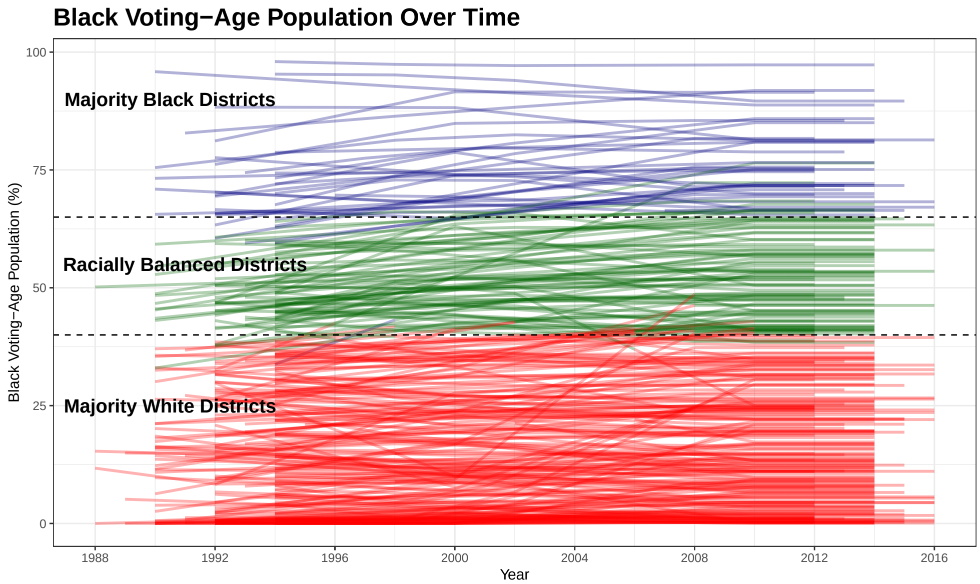

Figure B.2: **The changes in black voting-age population over time by racial regime**
Note: This graph visualizes the changes in black voting-age population for all districts over time as well as their corresponding racial regime. It demonstrates that very few districts experience a dramatic demographic change such that the percentage of black voting-age population crosses the boundaries of racial regime (i.e., 40% and 65%).

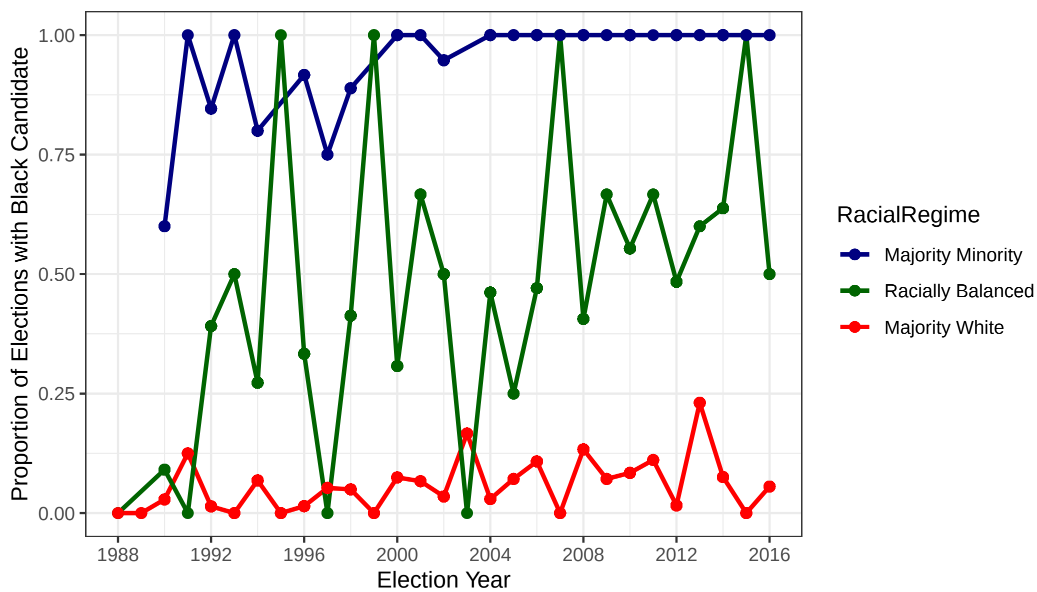

Figure B.3: **Overtime Variation of Black Candidate Emergence**
Note: This graph portrays the proportion of elections with black candidates over time by the racial regime.

B.4 Overtime Variation of the Racial Margin of Victory

To study whether there is any temporal pattern in the racial margin of victory, Figure B.4 plots the average values of the racial margin of victory over time. It also accounts for the racial regime. The figure illustrates that there is a clear temporal pattern for majority minority districts (black candidates have increasingly performed better relative to their white counterparts over time), whereas no such pattern can be observed for racially balanced and majority white districts.

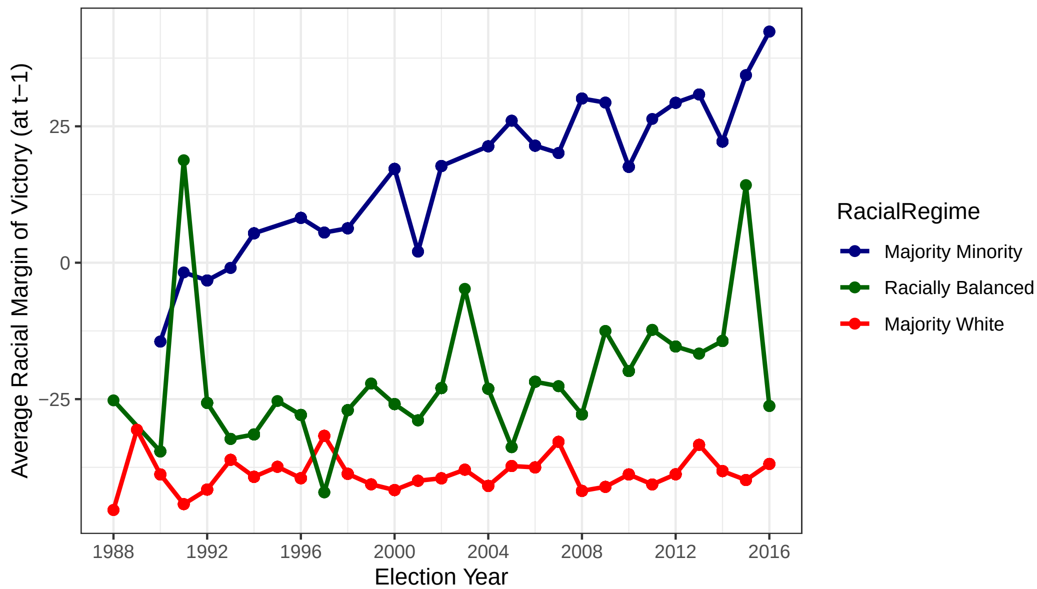

Figure B.4: **Overtime Variation of the Racial Margin of Victory**
Note: This graph portrays the proportion of elections with black candidates over time by the racial regime.

B.5 Registration by Race

In the State of Louisiana, the proportion of registered voters who identify themselves as neither white or black ranges from about 0.04 to 0.05 (Figure B.5). This implies that Louisiana mayoral elections are appropriate cases for validating the logical model based on Assumption 1.

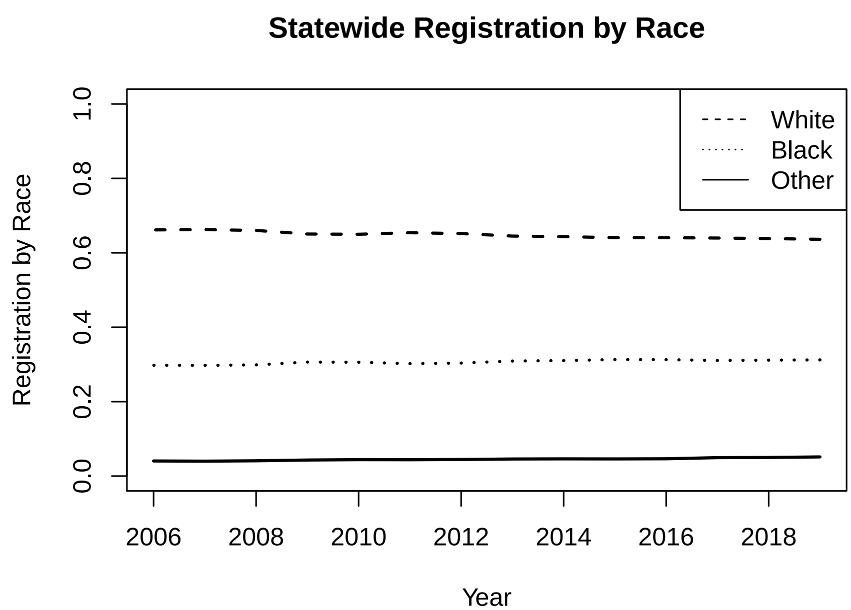

Figure B.5: **Registered Voters by Self-Reported Race**
Note: This plot represents the statewide voter registration record by self-reported race in Louisiana from 2006 to 2019. The data was collected from Louisiana Secretary of State website. It demonstrates that the proportion of registered voters who identify themselves neither as white or black ranges from about 0.0401 to 0.0515, giving some justification for considering non-partisan mayoral elections as biracial elections.

C Additional Findings

This section reports the results of additional analyses that support the claim in the article.

C.1 Count Model Regression Results

To buttress the claim that the racial margin of victory (and thus the logical model) does not predict the number of black candidates running for office, I run a series of count regressions (count models) with the number of black candidates as the dependent variable. Figure C.1 report the regression results. In Regressions 3-4, I dropped two observations which have 10 and 11 black candidates as outliers. I find that regardless of model specification the racial margin of victory has no statistically significant association with the number of black candidates (at the 0.05 level).

	Regression 1	Regression 2	Regression 3	Regression 4
Intercepts	-.238	-.238	-.293	-.293
	(.145)	(.145)	(.147)	(.147)
Racial Margin	.002	.002	.002	.002
	(.001)	(.001)	(.001)	(.001)
% blacks	.012	.012	.013	.013
	(.002)	(.002)	(.002)	(.002)
Specification	Poisson	Negative Binomial	Poisson	Negative Binomial
Outliers			dropped	dropped
N	526	526	524	524

Table C.1: **Regression Results**
Note: This table reports the results of count regressions.

C.2 Model Prediction with Placebo Outcomes

ePCP (win): 78.0 ePCP (female): 51.9

Cor(female, black) = 0.16

C.3 Extended Regression Results

C.4 Robustness Checks for Bayesian Logistic Regressions

To ensure the internal validity of the regression results, I performed multiple robustness checks by excluding all unopposed elections (column 4 in Table 2) as well as employing different cutoff points for racial regime and subsetting data before and after 2005 (see below). I apply these robustness checks since my original results could stem from a particular definition of racial regime, racial electoral performance based on unopposed elections, and data within particular time periods. Moreover, I conducted placebo tests using the presence of female candidates as a placebo outcome to check if the variable of interest only affects outcome values related to racial politics. The idea is that if the proposed model is theoretically sound, it must *not* predict female candidate emergence. Figure C.1 portrays the results of the placebo tests. The results demonstrate that both the black population size and the racial margin of victory do not have any bivariate relationship with female candidate emergence. One exception is that in majority minority districts there seems to be a slight association between the racial margin of victory and female candidate emergence. I suspect that this may be due to a potential correlation between race and gender of candidates. My substantive results are not susceptible to differences in measurements and the presence of unopposed elections, while they are not a product of mere chances.

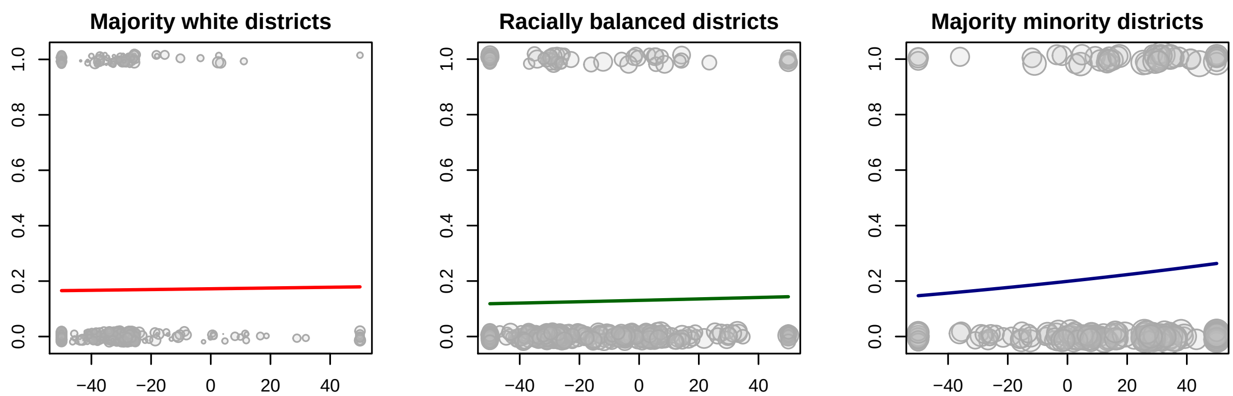

Figure C.1: **Placebo Tests with Female Candidate Emergence as the False Outcome Variable**
Note: This figure visualizes the results of placebo tests showing the bivariate relationships between racial electoral performance and the presence of female candidates. If the proposed model is theoretically sound, it must not predict female candidate emergence, and I confirmed this point in all racial regime. The size of open circles is proportional to the size of black voting-age population.

Below, I report multiple robustness checks by employing different cutoff points for racial regime, excluding all unopposed elections, and subsetting data before and after 2005. The estimated posteriors with lower and upper credible intervals are shown in Table C.2.

	40/65%	35/65%	30/65%	25/65%
REP (majority minority)	.039 (.025, .053)	.039 (.024, .054)	.042 (.025, .059)	.040 (.022, .058)
REP (racially balanced)	.051 (.039, .063)	.048 (.037, .060)	.047 (.036, .057)	.047 (.036, .057)
REP (majority white)	.062 (.039, .087)	.068 (.044, .094)	.067 (.042, .093)	.067 (.042, .092)
Covariates	✓	✓	✓	✓
RE (Years)	✓	✓	✓	✓
RE (Municipalities)	✓	✓	✓	✓
N	2037	2037	2037	2037

	40/70%	40/75%	40/80%	35/70%
REP (majority minority)	.039 (.026, .053)	.039 (.025, .052)	.040 (.027, .054)	.040 (.026, .055)
REP (racially balanced)	.051 (.040, .063)	.051 (.040, .063)	.053 (.042, .064)	.050 (.039, .060)
REP (majority white)	.075 (.039, .113)	.085 (.041, .132)	.038 (-.028, .110)	.073 (.037, .111)
Covariates	✓	✓	✓	✓
RE (Years)	✓	✓	✓	✓
RE (Municipalities)	✓	✓	✓	✓
N	2037	2037	2037	2037

	30/75%	25/80%	Before 2005	After 2005
REP (majority minority)	.043 (.026, .060)	.043 (.026, .061)	.043 (.021, .065)	.037 (.016, .057)
REP (racially balanced)	.048 (.038, .058)	.049 (.040, .060)	.049 (.032, .066)	.061 (.043, .081)
REP (majority white)	.082 (.038, .129)	.036 (-.027, .106)	.059 (.031, 087)	.086 (.006, .201)
Covariates	✓	✓		
RE (Years)	✓	✓	✓	✓
RE (Municipalities)	✓	✓	✓	✓
N	2037	2037	1166	871

Table C.2: **Estimated Results with Different Cutoff Points and Subset of Data**
Note: This table shows the posterior estimates of the effect of racial electoral performance on minority candidate emergence using different cutoff points for racial regime and subsets of data. The results demonstrate that the original result is not susceptible to these changes in the definition and time periods for analysis.