Appendix B of Online Appendix

For "A Logical Model for Predicting Minority Representation: Application to Redistricting and Voting Rights Cases"

Yuki Atsusaka

Table of Contents

A Derivation of the Logical Model **1**

B Additional Findings and Discussions **2**

B.1 Previous Studies Predicting Minority Representation 2

B.2 Registration by Race in Louisiana 3

B.3 Results for Extended Regressions with 34 Additional Variables 3

B.4 Analyses for Internal Validity 5

B.5 Minority Retreat v. Hopeless Entry 8

B.6 Predicting the Number of Minority Officeholders 9

A Derivation of the Logical Model

See the main Online Appendix for this section.

B Additional Findings and Discussions

This section summarizes the results of additional analyses that supplement the main analyses reported in the main text.

B.1 Previous Studies Predicting Minority Representation

Table B.1 summarizes previous studies that predict district-level majority candidate emergence and electoral success. Here, I limit the scope of my literature review to those studies concerning minority representation *within the same electoral systems* reflecting the purpose of the logical model. For a rich body of studies concerning the effect of electoral systems, especially converging from at-large elections to single-member districts, on minority descriptive representation, see Abott and Magazinnik (2020). I leave future research to investigate how the logical model can be used to predict the impact of changing electoral systems on minority descriptive representation.

The table shows that previous studies have used various statistical methods and data when reporting the predicted probability of minority candidate emergence and electoral success. One important implication of this diversity in methodologies and data for this article is that it is challenging to *generalize* such predicted probabilities to other contexts outside of the particular studies. In contrast, the logical model applies to any district because its functional form does not depend on any statistical method and data.

	Statistical Method	**Data/Context**
Fraga, Juenke and Shah (2020)	nonparametric regression	state legislature
Lublin et al. (2019)	logit	state legislature, U.S. House
Hicks et al. (2018)	logit	state legislature
Shah (2017)	multinomial logit	city
Juenke and Shah (2016)	logit	state legislature
Juenke and Shah (2015)	probit	state legislature
Shah (2014)	Heckman's selection model	Louisiana municipality
Juenke (2014)	probit	state legislature
Casellas (2011)	probit	state legislature, U.S. House
Marschall, Ruhil and Shah (2010)	Mullahy's hurdle Poisson model	city, school district
Branton (2009)	zero-inflated Poisson regression	U.S. House (primary)
Lublin et al. (2009)	logit	state legislature
Marschall and Ruhil (2006)	probit (pooled & random effects)	mayor
Burns (2003)	OLS	Southern city
Tate (2003)	descriptive statistics	U.S. House
Canon (1999)	generalized event count model	U.S. House
Epstein and O'Halloran (1999)	multinomial logit	U.S. House
Lublin (1999)	logit	U.S. House
Lublin (1997)	logit	U.S. House
Cameron, Epstein and O'halloran (1996)	multinomial logit	U.S. House

Table B.1: **Various Methods and Data in Predicting Minority Representation**
Note: This table illustrates the variety of statistical methods and data in selected previous studies that predict district-level minority candidate emergence and electoral success.

B.2 Registration by Race in Louisiana

In the State of Louisiana, the proportion of registered voters who identify themselves as neither white nor black ranges from about 0.04 to 0.05 (Figure B.1). This implies that Louisiana mayoral elections are appropriate cases for validating the logical model based on Assumption 1 (Biracial Elections).

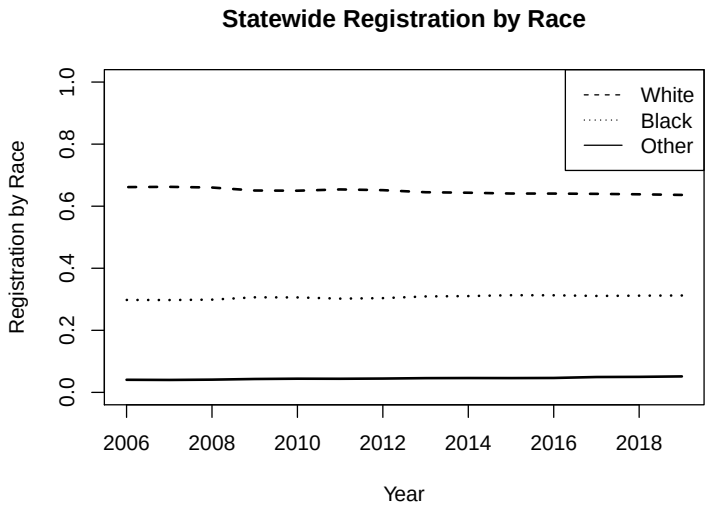

Figure B.1: **Registered Voters by Self-Reported Race Over Time**
Note: This plot represents the statewide voter registration record by self-reported race in Louisiana from 2006 to 2019. The data was collected from Louisiana Secretary of State website. It demonstrates that the proportion of registered voters who identify themselves neither as white or black ranges from about 0.0401 to 0.0515, giving a justification for considering non-partisan mayoral elections as biracial elections.

B.3 Results for Extended Regressions with 34 Additional Variables

To highlight the relative predictive power of the logical model to *ad hoc* regressions, I compare the ePCPs obtained from the logical model with those from LPM and logistic regressions with 34 additional variables (which previous research tends to include) using the Louisiana mayoral election data. These variables include the percentage of blacks and whites with B.A., a dummy variable for the election cycle, a dummy variable for open races, the percentage of whites who are 65 and over, a measure for human density, dummy variables for election years, and dummy variables for municipalities.

Table B.2 shows that the logical model performs as good as and even better than these heavily parameterized regressions. The results are rather remarkable because despite that these extended regressions are meant to have high *in-sample* ePCPs and fit (and even overfit) individual data points, the logical model still dominates them in terms of predictive power.

	Minority Candidate Emergence			Minority Electoral Success		
	Logical Model	LPM	Logit	Logical Model	LPM	Logit
All Sample (N=2037)	88.6	84.3	87	93.5	88.2	90.6
0<C<40 (1307)	94.1	90.6	91.6	97.8	96.3	97
40<C<65 (494)	72.7	73.6	75	86	80.5	82.7
65<C<100 (236)	91	93	100	85.9	80.8	81.9
Urban (57)	66.7	95.1	100	91.2	91.8	100
Suburban (715)	89.9	85.6	87.8	94.9	90.7	93.8
Rural (1258)	88.9	84.7	87.8	92.9	87.8	90
Open-Seat (584)	84.4	82.2	85	91.3	85.7	88.7
Not Open-Seat (1453)	90.2	85.9	89	94.4	89.9	92.7
Uncontested Elections (870)	97.3	94.1	96.5	97.6	94.6	96.7
Contested Elections (1167)	82	79.1	81.1	90.5	84.2	87.3
On-Cycle (1805)	88.7	84.5	87	93.9	88.5	90.9
Off-Cycle (232)	87.1	85.1	88.1	90.5	87.1	88.6
At-Large (Councils) (880)	93.5	90.2	92.9	94.7	90.7	92.9
District (Councils) (300)	78	78	81.7	88	82.9	85.8
Mixed (Councils) (857)	87.2	82.9	85.9	94.3	89.2	92.7
Before 1994 (195)	91.1	86.2	90.3	94.9	92.1	94.9
After 1994 (1842)	88.3	84.2	86.7	93.4	87.9	90.2

Table B.2: **Predictive Performance of the Logical Model and Extended Regressions**
Note: This table reports the predictive performance of the logical model and linear and logistic regressions by (in-sample) ePCPs.

B.4 Analyses for Internal Validity

To examine the internal validity of my argument, this subsection examines several key observable implications from the logical model. Such observable implications include: (1) the majority of elections with minority candidate emergence features 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, (4) the logical model does not predict woman candidate emergence and electoral success (assuming that being a women candidate is not correlated with being a black candidate), and (5) the most recent election contains the most updated and thus valuable information about future electoral results. I found evidence for all of these expectations using the Louisiana mayoral election data.

B.4.1 Majority of Elections with Minority Candidates Only Features One Minority Candidate

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. This statistic indicates that black candidates more or less coordinated well among themselves to avoid vote splitting by black voters.

B.4.2 The Logical Model Does Not Predict the Number of Minority Candidates

Second, I found that the logical model does not predict the number of black candidates. The left panel of Figure B.2 plots the number of black candidates on ballots over the model predictions as a contour plot. It shows that there is no clear relationship between the number of black candidates and the racial margin of victory (x-axis). To provide 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. Table B.3 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). 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.

	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 B.3: **Results for Count Regressions**
Note: This table reports the results of count regressions. Regardless of model specification, the racial margin of victory does not have a statistically significant association with the number of black candidates.

B.4.3 The Number of Minority Candidates Follows A Random Pattern

Third, I found that the number of black candidates follows a "random pattern" (instead of model predictions) and, more specifically, a power law in its tail.[1] The power law can be observed in diverse random phenomena where many observations are clustered around some typical values (e.g., the intensity of wars, the severity of terrorist attacks, and the frequency of US family names) (Clauset, Shalizi and Newman, 2009). When empirical data follows the power law (distribution), very few values are observed with significantly high frequencies with many other values with low frequencies. The right plot of Figure B.2 visualizes this pattern for the number of black candidates in our data. The hypothesis test suggests that the null hypothesis that the data follow a power-law distribution cannot be rejected (p-value=0.79).

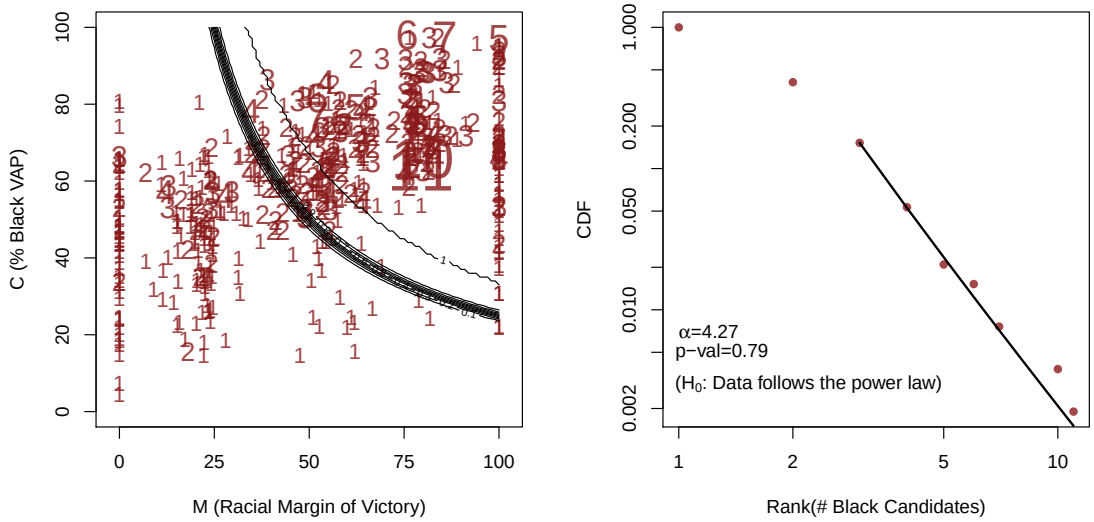

Figure B.2: **The Model Does Not Predict the Number of Black Candidates**
Note: The left panel plots the number of black candidates running for office over the parameter space of the logical model with model predictions as a contour plot. The right panel shows that the number of black candidates follows a power law distribution.

B.4.4 The Logical Model Predicts Minority Representation, but Not Women's Representation

Fourth, I found that the logical model does *not* predict female candidate emergence and electoral success as two placebo outcomes. Table B.4 suggest that the logical model only predicts woman candidate emergence with only about 50% and electoral success with about 75% accuracy on average (via ePCP). This indicates that the model is indeed a logical model for racial *minority* representation and not other types of representation.

[1] 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).

	Woman Candidate Emergence			Woman Electoral Success		
	Logical Model	LPM	Logit	Logical Model	LPM	Logit
All Sample (N=2037)	48.7	53.5	53.5	74.7	72.9	72.9
0<C<40 (1307)	43.5	52.8	52.9	82.8	72.3	72.3
40<C<65 (494)	52.1	52.9	53	73.9	78.7	78.7
65<C<100 (236)	70.3	61.3	61.3	31.3	65.7	65.7
Urban (57)	40.4	66.2	66.6	71.9	90.1	90.1
Suburban (715)	44.7	56.3	56.4	78.6	73.6	73.6
Rural (1258)	51.5	52.5	52.5	72.7	72.3	72.3
Open-Seat (584)	40.3	57.8	57.8	74.8	68.3	68.3
Not Open-Seat (1453)	52	52.8	52.9	74.6	75	75
Uncontested Elections (870)	51.9	51.2	51.2	77.8	71.8	71.8
Contested Elections (1167)	46.3	56.5	56.6	72.4	74.1	74.1
On-Cycle (1805)	49.8	53.3	53.4	75	73.3	73.3
Off-Cycle (232)	40.1	57.7	57.8	72	70	69.9
At-Large (Councils) (880)	45	53.4	53.4	68.3	64	64
District (Councils) (300)	43.1	59.1	59.3	74.7	90	90
Mixed (Councils) (857)	54.4	53.5	53.6	81.2	78.5	78.5
Before 1994 (195)	69	58.6	58.7	83.3	80.3	80.4
After 1994 (1842)	46.5	54.1	54.2	73.8	72.2	72.2

Table B.4: **Predicting Women Representation as Placebo Outcomes**
Note: This table reports the predictive performance of the logical model and linear and logistic regressions by (in-sample) ePCPs.

B.4.5 The Most Recent Election Contains the Highest-Quality Information about Electoral Results in the Upcoming Election

Fifth, readers may wonder what if we use the racial margin of victory for the last two elections or even three elections, instead of the immediate last election. Does using the last two and three elections contribute to the higher predictive performance of the model? To answer this question, I compute model predictions based on two *modified* racial margin of victory terms which are computed by taking an arithmetic mean of the racial margin of victory from the last two and three elections, respectively. The results are shown in Table B.5 (the number of observations was reduced to 1431 in total as the modified racial margin of victories are not available for all elections in the data). It demonstrates that these extended models yield a poorer predictive performance than the original logical model in all subsets of data. This result further supports the internal validity of the logical model: results of the most recent election contain the most updated (and thus most valuable) information about the expected behavior of minority and white voters (and thus future vote shares). From the practical point of view, this is highly advantageous because users of the logical model can predict minority candidate emergence and electoral success with the minimal amount of information (i.e., they only need to collect information for M from time **T-1** instead of at time **T-1**, **T-2**, and even **T-3**).

	Minority Candidate Emergence			Minority Electoral Success		
	M_{t-1}	$\frac{1}{2}(M_{t-1}+M_{t-2})$	$\frac{1}{3}(M_{t-1}+M_{t-2}+M_{t-3})$	M_{t-1}	$\frac{1}{2}(M_{t-1}+M_{t-2})$	$\frac{1}{3}(M_{t-1}+M_{t-2}+M_{t-3})$
All Districts (1431)	87.9	84.1	83.7	92.7	79.6	79.4
0<C<40 (905)	93.2	91	91.1	97.1	93.2	93.7
40<C<65 (350)	72	59.7	58.4	83.5	39.8	37.6
65<C<100 (176)	92.6	96.6	96.6	88.4	89.2	89.2
Urban (39)	71.8	87.3	87.2	92.3	66.8	66.7
Suburban (507)	89.6	85.6	85.5	94.3	86	85.6
Rural (880)	87.7	83	82.6	91.8	76.4	76.3
Open-Seat (422)	83.8	85.5	84.9	89.6	80.6	80.8
Not Open-Seat (1009)	89.6	83.5	83.2	94	79.3	78.8
Uncontested (606)	96.9	86.9	85.8	97.4	86.8	85.6
Contested (825)	81.3	81.9	82.3	89.3	74.4	74.8
On-Cycle (1261)	87.9	83.3	83	93.1	78.9	78.7
Off-Cycle (170)	87.6	89.7	89.1	90	85	84.4
At-Large (Councils) (615)	92.5	90	89.5	94.2	88.2	88.1
District (Councils) (213)	77.1	71.7	70.4	87.2	57.6	56
Mixed (Councils) (603)	87.1	82.3	82.6	93.1	78.7	78.8
Before 1994 (1)	100	100	100	100	100	100
After 1994 (1430)	87.9	84	83.7	92.7	79.6	79.4

Table B.5: **Computing M from Multiple Past Elections.**
Note: This table reports the predictive performance of the logical model with different ways of computing M in ePCPs. The original operationalization of M (Columns 1 and 4) yields higher prediction performance than its modified versions.

B.5 Minority Retreat v. Hopeless Entry

The empirical evaluation of the logical model's predictions reveals two types of elections for which the logical model fails to correctly predict minority candidate emergence. First, there are several cases in which minority candidates did not emerge even though the model predicts that they have a high chance of winning. Second, there are several elections in which minority candidates emerged even though the model predicts a very low probability of winning. Here, I label the first cases as *minority retreat* and the second situations as *minority hopeless entry*. Recently, the literature on minority representation has extensively studied the minority retreat problem (Shah, 2014; Juenke, 2014; Juenke and Shah, 2015, 2016; Fraga, Juenke and Shah, 2020), while the minority hopeless entry has been less systematically examined. In this subsection, I investigate the relative frequency of the two problems in our data to understand for which types of cases "unpredictable candidate emergence" occurs.

Table B.6 reports the main findings, illustrating that minority candidates are more likely to emerge in elections despite the low odds of winning than retreat from districts despite the high probability of winning. Columns 1-2 report the relative frequencies (in percentage) of elections in which minority candidates did not run while the logical model predicts that there is **0.5 or higher** probability of winning using the two different data sets. Using 0.5 as a cutoff point offers the most liberal test (a test under the most favorable condition for the hypothesis) for the minority retreat hypothesis since it may include more cases where minority candidates could potentially retreat. These results suggest that while we observe some cases where minority candidates "failed to run" despite a high chance of winning, **such cases do not consist the majority of electoral contests**. Indeed, they imply that most minority candidates properly assess their chance of winning

and enter electoral races when they see a viable opportunity to secure seats even outside majority-minority districts (%minority < 50), which is more consistent with the argument made by Grofman, Handley and Lublin (2001) and Lublin et al. (2019).

Columns 3-4 report the relative frequencies (in percentage) of districts in which minority candidates emerge while the logical model predicts that there is **0.1 or lower** probability of winning. They show that there are consistently more cases that fit the minority hopeless entry than the minority retreat (one exception is majority-white districts in the Louisiana data, but its small number of cases prevents us from drawing a more definitive conclusion). As discussed in the main text, this may represent the fact that minority candidates sometimes run for office due to non-instrumental or long-term rational reasons.

	Minority Retreat		**Minority Hopeless Entry**	
	Louisiana Data	State Legislative Data	Louisiana Data	State Legislative Data
All Districts	3.2% (10/312)	1.9% (6/320)	12.4% (83/188)	10.5% (103/972)
C<50	11.1% (2/18)	0% (0/11)	8.5% (129/1542)	7.8% (70/902)
C≥50	2.7% (8/294)	2.1% (6/289)	44.2% (83/188)	47.1% (44/76)

Table B.6: **Frequency of Minority Retreat and Hopeless Entry**
Note: This table shows the relative frequency (in percentage) of elections in which minority candidates did not appear or emerged again the logical model's predictions. Specifically, Columns 1-2 show the frequency of elections in which minority candidates did not appear even though the model predicts that there is 0.5 or higher probability of wining (i.e., elections from which minority candidates retreated again the high odds of winning). Columns 3-4 report the frequency of elections in which minority candidates run for office even though the logical model predicts that there is 0.1 or lower probability of winning (i.e., elections minority candidates entered despite the low odds of winning).

B.6 Predicting the Number of Minority Officeholders

Figure B.3 and Figure B.4 demonstrate that the logical model can also accurately predict the number of minority officeholders in an entire jurisdiction, such as city, county, and state. Each histogram summarizes the predicted number of minority officeholders obtained from Monte Carlo simulations based on the logical model (see below) for each year or state. The true number of minority officeholders is shown in a dashed line. To evaluate the prediction accuracy, the modal gap (the difference between the modal prediction and the truth) is reported for each panel. For most panels (24/29 in Figure B.3 and 32/36 in Figure B.4), the modal gap is 0 or between -2 and 2. These results suggest that not only can the logical model predict district-level minority representation, but it can also accurately predict *jurisdiction-level* minority representation.

Now I present how the logical model can be used to predict the number of minority officeholders. Let **M**=$(M_1, M_2, ..., M_D)$ be a vector of racial margin of victories for $i = 1, ..., D$ districts. Let **C**=$(C_1, C_2, ..., C_D)$ be a vector of percentages of minority voters for $i = 1, ..., D$ districts. The first step is to predict the probability of minority electoral success for each district and create a vector of predicted probabilities. I denote such a probability vector by **P**=$(P_1, P_2, ..., P_D)$. The second step is to perform Monte Carlo simulations in which we repeat sampling from a Binomial distribution with a mean vector of **P** for 1000 times. Or equivalently, we draw a sample of 0 (minority failure) or 1 (minority success) from a Bernoulli distribution with a "success probability" (as appropriately termed both for the logical model and distribution) $P_1, P_2, ..., P_D$ (for each district), respectively and we repeat this process 1000 times. The final step is to summarize the distribution of repeated samples via histogram. A great advantage of this approach is that the final result reflects the uncertainty around the model prediction for each district.

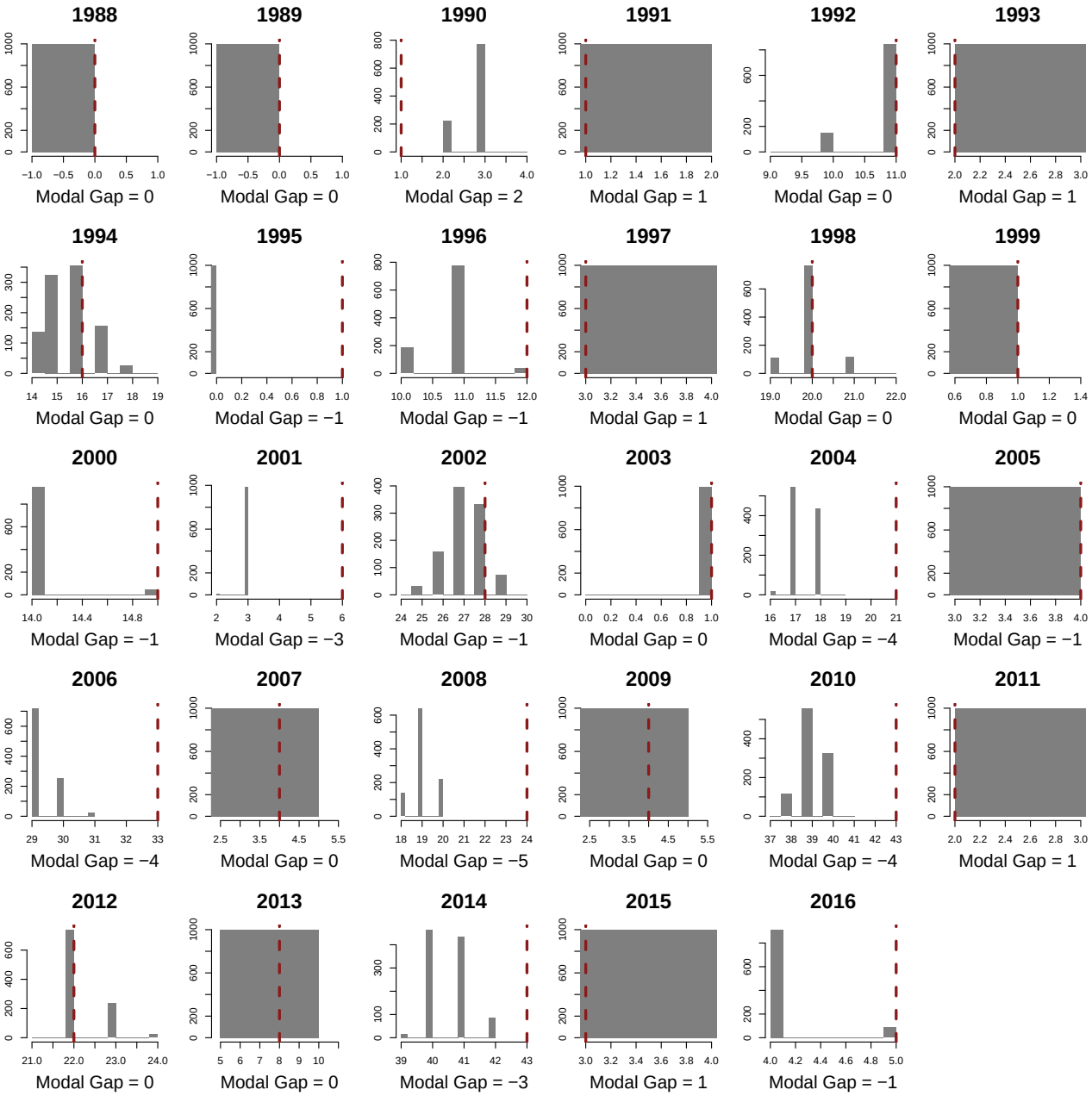

Figure B.3: **Predicted Number of Black Mayors in Louisiana**
Note: Each histogram summarizes the predicted number of minority officeholders obtained from Monte Carlo simulations based on the logical model for each year. The true number is shown in a dashed line. The modal gap (the difference between the modal prediction and the truth) is reported for each year.

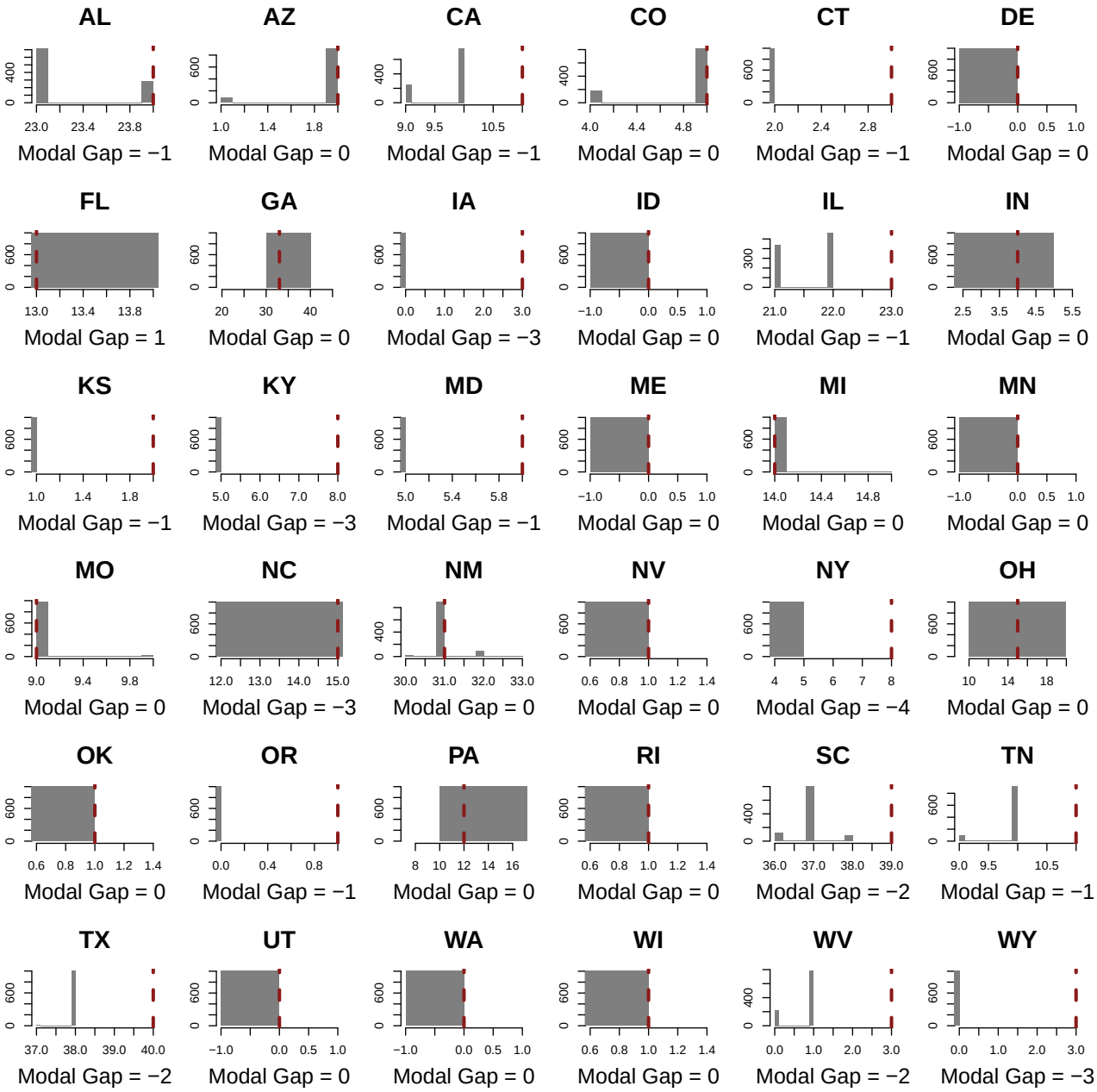

Figure B.4: **Predicted Number of Minority State Legislators in 36 States**
Note: Each histogram summarizes the predicted number of minority officeholders obtained from Monte Carlo simulations based on the logical model for each state. The true number is shown in a dashed line. The modal gap (the difference between the modal prediction and the truth) is reported for each state.

References for Online Appendix

Abott, Carolyn and Asya Magazinnik. 2020. "At-Large Elections and Minority Representation in Local Government." *American Journal of Political Science* .

Branton, Regina P. 2009. "The importance of race and ethnicity in congressional primary elections." *Political Research Quarterly* 62(3):459–473.

Burns, Peter F. 2003. "The South, Medium-Sized Cities, and a New Look at the Determinants of African-American Representation." *Politics & Policy* 31(4):620–647.

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. 2011. *Latino representation in state houses and congress*. Cambridge University Press.

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

Epstein, David and Sharyn O'Halloran. 1999. "A social science approach to race, redistricting, and representation." *American Political Science Review* pp. 187–191.

Fraga, Bernard L, Eric Gonzalez Juenke and Paru Shah. 2020. "One Run Leads to Another: Minority Incumbents and the Emergence of Lower Ticket Minority Candidates." *The Journal of Politics* 82(2):771–775.

Grofman, Bernard, Lisa Handley and David Lublin. 2001. "Drawing effective miority districts: A conceptual framework and some empirical evidence." *NCL rev.* 79:1383.

Hicks, William D, Carl E Klarner, Seth C McKee and Daniel A Smith. 2018. "Revisiting majority-minority districts and black representation." *Political Research Quarterly* 71(2):408–423.

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.

Lublin, David. 1997. *The paradox of representation: Racial gerrymandering and minority interests in Congress*. Princeton University Press.

Lublin, David. 1999. "Racial redistricting and African-American representation: A critique of "do majority-minority districts maximize substantive black representation in Congress?"." *American Political Science Review* 93(1):183–186.

Lublin, David, Lisa Handley, Thomas L Brunell and Bernard Grofman. 2019. "Minority success in non-majority minority districts: Finding the "sweet spot"." *Journal of Race, Ethnicity and Politics* 5(2):275–298.

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.

Marschall, Melissa J and Anirudh VS Ruhil. 2006. "The pomp of power: Black mayoralties in urban America." *Social Science Quarterly* 87(4):828–850.

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.

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.

Tate, Katherine. 2003. *Black faces in the mirror: African Americans and their representatives in the US Congress*. Princeton University Press.