# Online Appendix for: Evaluating the Minority Candidate Penalty with a Regression Discontinuity Approach

## A1 Description of Elections Used in the Analysis

We use data on state legislative candidates and outcomes in our regression discontinuity analysis. As noted in the main text, we begin with information on primary election outcomes where the race of the top vote winners differed, where the winning margin serves as our RD's running variable. In systems that hold runoff elections to choose between the top two candidates in a multi-candidate primary, we use the runoff results to calculate the winning margin, rather than the first round. Some of the 42 states we examine have multi-member state legislative districts or use primary systems where multiple candidates advance to the general election from a single primary ballot. We exclude systems with non-partisan "top two" primary systems that advance the two highest vote-getters regardless of party, as they do not produce party nominees. In the case of multi-member districts, we focus on the "last in" and "first out" candidates when calculating the primary win margin, comparing the candidate who most narrowly won nomination to the person who most narrowly lost it.

Table A1 reports the total number of multiracial primary elections we use in the analysis, excluding uncontested general elections as well as races omitted due to missingness from other sources[1], by year and party; there are more Democratic primaries than Republican ones, but both parties are well-represented across both years of the dataset. Furthermore, in both 2018 and 2020 the breakdown is about three-quarters lower-chamber elections (State House) and one-quarter upper (State Senate).

[1] These are mainly cases where the running variable, "minority candidate primary win margin (over white candidate)," is missing because none of the top candidates in a contested multiracial primary are white.

Table A1: Election counts

| | Democratic | Republican |
|---|---|---|
| 2018 | 142 | 53 |
| 2019 | 6 | 0 |
| 2020 | 153 | 46 |

We note that the districts included in the RD sample are not exclusively “majority-minority districts" (fewer than one-quarter have a white population under 50%), nor are they overwhelmingly safe Democratic districts. Figure A1 explores the partisan composition of the districts in our sample by plotting the density of Democratic two-party presidential voteshare in 2016 for the state legislative districts in the RDD sample, and for the rest of the US. Presidential voteshare data used here come from MIT's Election and Data Science Lab (they rely on aggregating precinct-level voteshare data to the level of the state legislative district) and have some missingness, such that about 9 in 10 of the districts in our sample are successfully matched to district presidential-vote data. Figure A1 indicates that districts in our sample are slightly more Democratic than average, but are generally not “safe" Democratic districts in the sense of having very high Democratic vote shares. Rather, the most common values for both our sample of districts and for the rest of the US fall within the competitive zone of 40-60% Democratic voteshare.

In addition to the distribution of districts included, some readers may wonder whether the effects differ across parties: could it be that Democrats (the majority of the sample) do not face an electoral penalty when nominating a minority candidate, but that Republicans do? We note the limitations of heterogeneity analysis; given the size of the sample used for the main study, any subgroup analysis is going to be somewhat underpowered. But we present a version of the main table here that restricts the sample to the 120 Republican primaries included in the main dataset. The estimates included in Table A2 are noisy and variable, but they do not suggest that Republicans face a substantial electoral penalty.

Figure A1: Comparing RD sample districts to the rest of the US

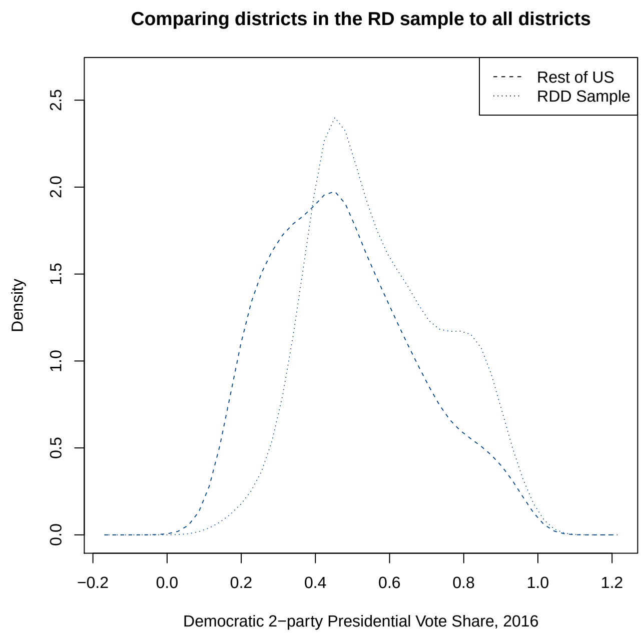



Table A2: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, Republican primaries only

| | | | |
|---|---|---|---|
| Minority nominee | 0.23 | 0 | 0 |
| | (0.08,0.46) | (-0.06,0.08) | (-0.11,0.08) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 28 | 28 | 41 |
| Bandwidth | 0.13 | 0.13 | 0.2 |

## A2 Alternative Treatment of Uncontested Elections

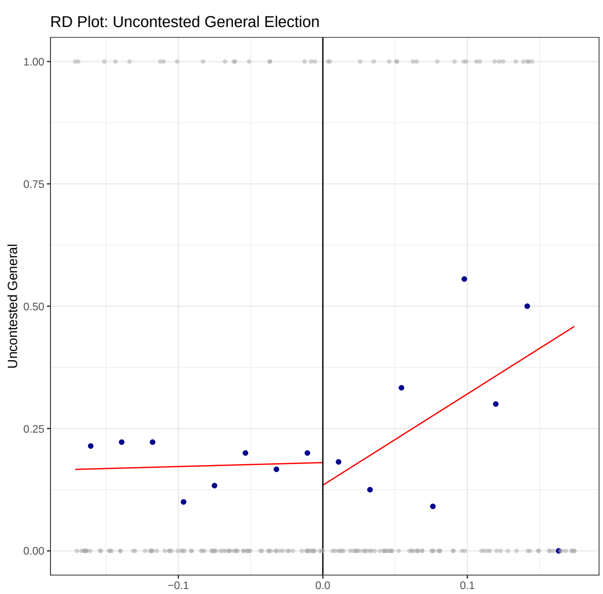
Our main specifications omit uncontested elections, considering an election as uncontested if one of the two major parties does not have a candidate on the general election ballot. These uncontested general elections are not informative about the question asked in the paper. In most cases, the opposing party had no nominee for the general election even *before* the primaries occurred, so there is no logical way that one party's nomination of a white candidate could change the presence of the other party's candidate on the ballot.

Nevertheless, in this section we explore an alternative approach to uncontested elections, setting the two-party vote share to 100% for winning parties (rather than missing) and to 0% for non-contesting parties in cases of uncontested general elections. This approach adds additional observations to the sample, but also introduces noise (as we are simply imputing votecounts in elections that did not occur). Table A3 reproduces Table 1 from the main paper using this alternative outcome measure. The specifications vary in size and direction, but still do not clearly point to a substantial electoral penalty. Further, Figure A2 below illustrates that the rate of uncontested elections is smooth across the cutpoint of minority-candidate win margin.

Table A3: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, with alternative voteshare measure

| | | | |
|---|---|---|---|
| Minority nominee | 0.05 | -0.02 | -0.03 |
| | (-0.11,0.21) | (-0.11,0.06) | (-0.13,0.07) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 212 | 195 | 379 |
| Bandwidth | 0.21 | 0.19 | 0.44 |

Figure A2: Examining smoothness of uncontested election status across the cutpoint

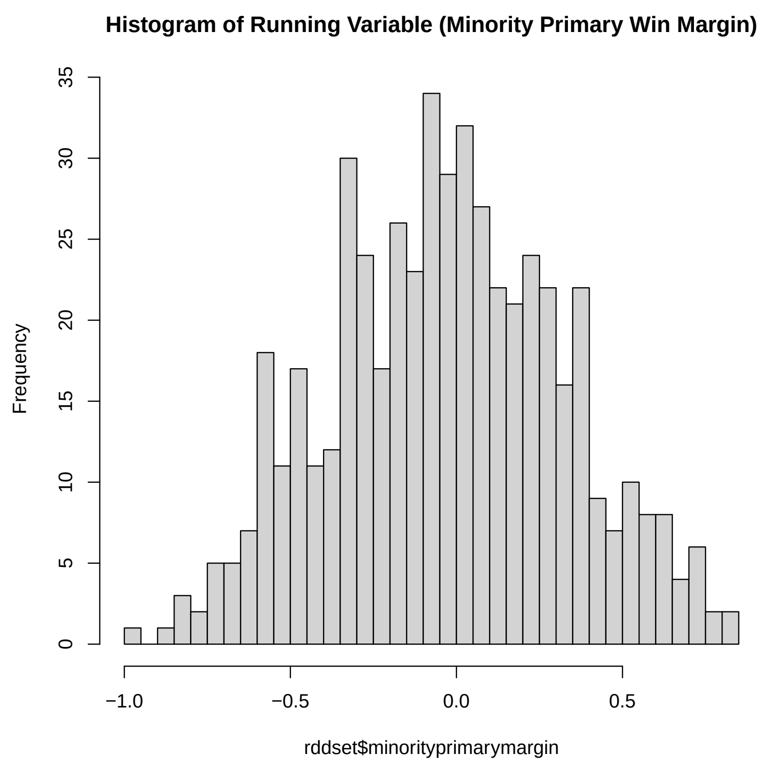


## A3 RDD Validity

In this section, we test the validity of the regression-discontinuity assumptions and explore robustness to alternate specifications.

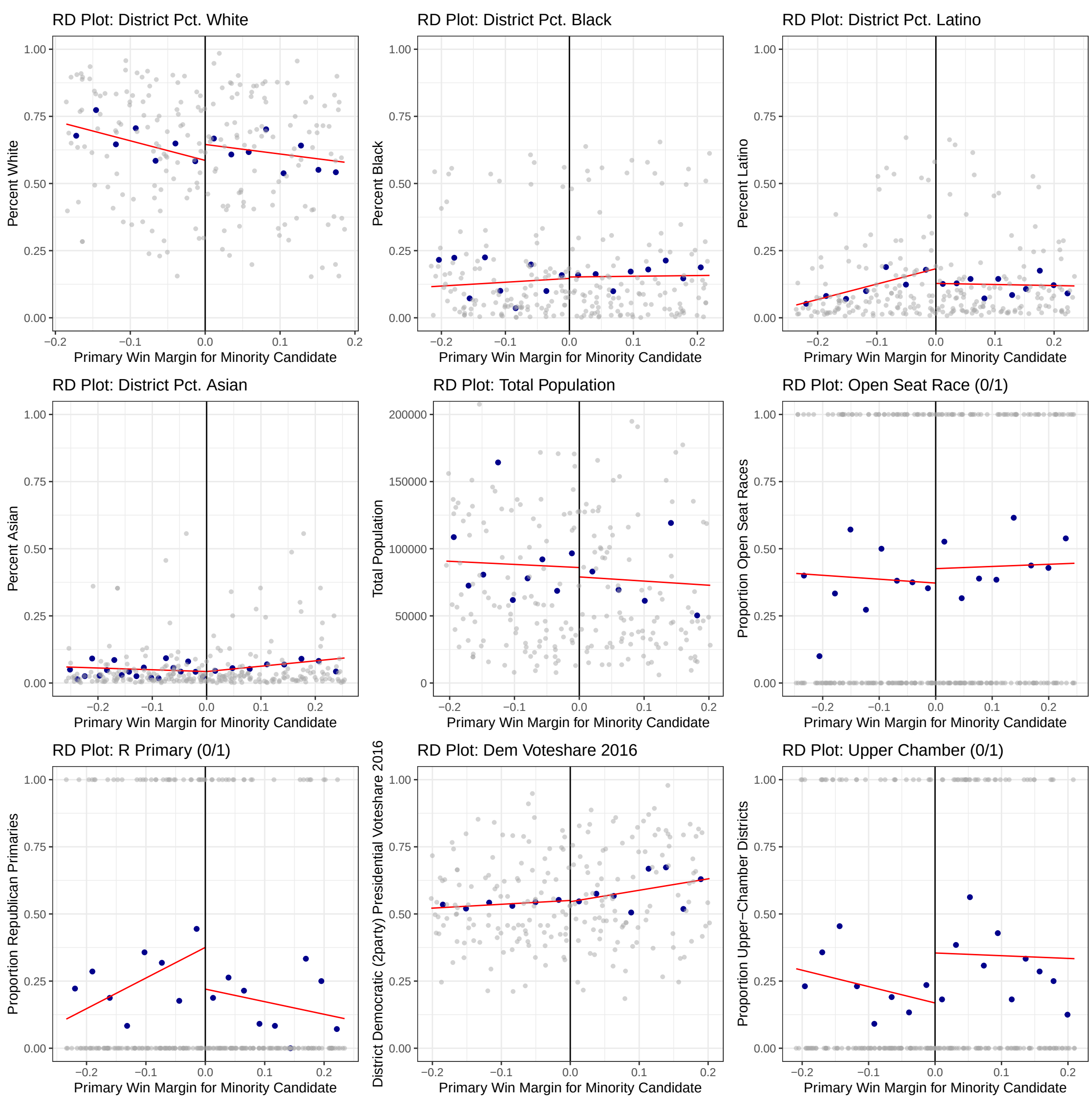
We begin by looking for evidence of sorting around the cutpoint. If units were able to select which side of the cutpoint they landed on, the RD setup would not be valid. We present a simple histogram of the running variable (Figure A3), which doesn't show any notable lopsidedness around the cutpoint. We also ran a test for continuity of the density functions for control and treatment units around the cutoff using the **rddensity()** package in R (Cattaneo, Jansson and Ma 2021), which failed to reject the null hypothesis of no manipulation at the cutoff.

Figure A3: Looking for sorting around the cutpoint




We next turn to background covariates, which ought to look smooth across the cutpoint: districts where minority candidates narrowly win the primary should not look systematically different from districts where they lose. To illustrate this smoothness, we graphically display the equivalent of our main RDD specification (local linear with "mserd" automated bandwidth selection), except that in this case we use a variety of pre-election covariates as the outcome variable. These can be thought of as a sort of placebo test: a significant "treatment effect" on these pre-election covariates would indicate a problem with the RD setup. We observe no such pattern in Figure A4: discontinuities at the cutpoint are relatively small, and none are significant (p-values range from .23 to .99).

Figure A4: Examining smoothness of background covariates across the cutpoint

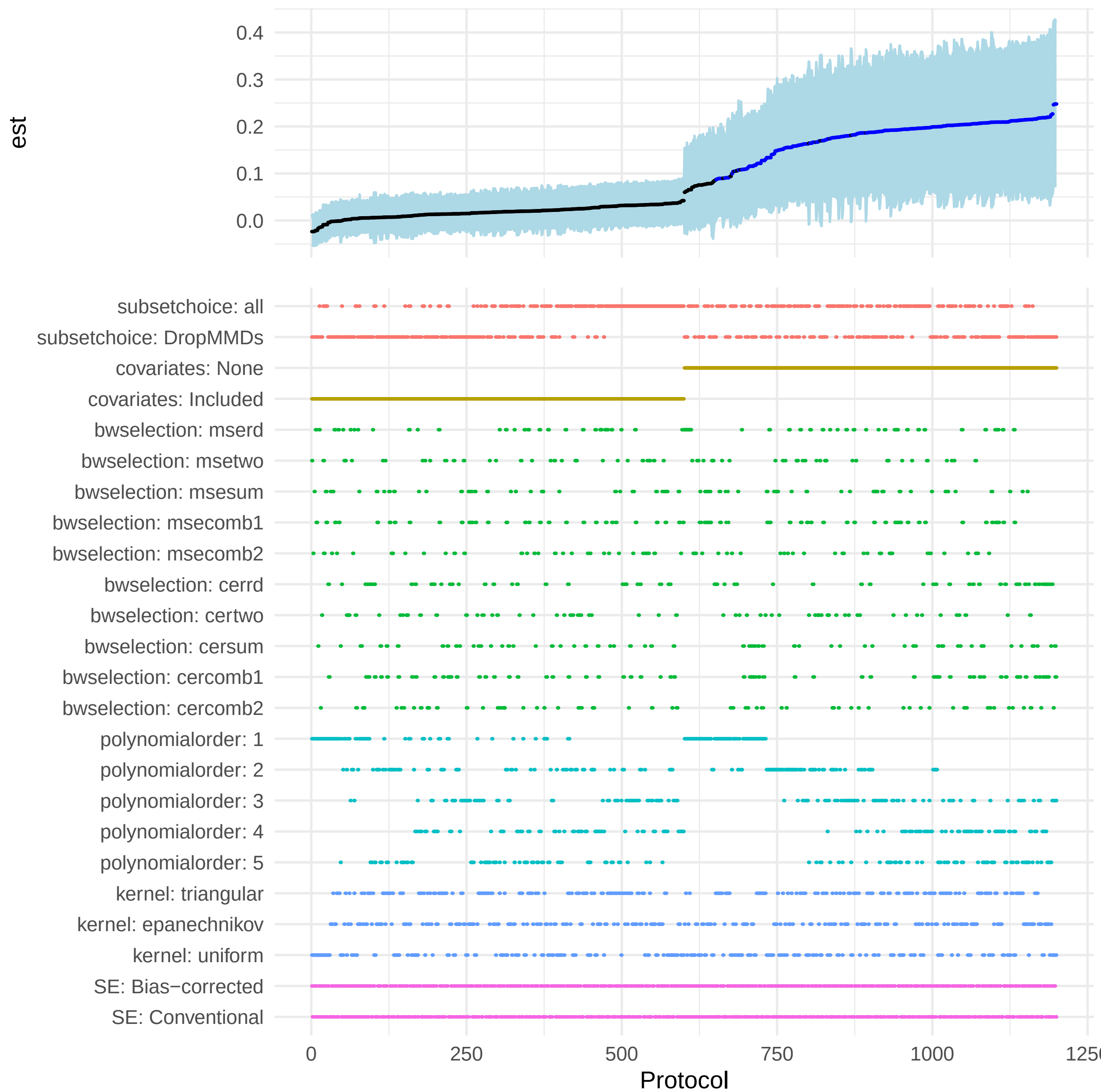


Finally, we present a plot that explores many plausible specifications for the main RD estimates presented in the paper. In estimating the electoral penalty/benefit to a party that nominates a minority candidate, our main approach in the paper uses all "factory default" settings from the rdrobust() package: a local linear specification with automated bandwidth selection using the "mserd" option and bias-corrected estimates with robust standard errors. In this plot, we begin to vary some of the data-management and analytic decisions made in those main estimates in order to explore how much those choices mattered for the estimates presented. The plot presents RD estimates from 1,200 specifications that vary these decisions. The top part of the plot displays the distribution of point estimates from those models along with 95% confidence intervals. The bottom section of the plot indicates which specification choices are associated with which estimates.

The choices vary as follows. First, the plot varies the data included ("subsetchoice") to rely either on the full analysis dataset from the main paper, or to omit the small number of multi-member districts that were included in our dataset and focus on single-member districts. We include specifications that do not include covariates, as well as specifications that incorporate available covariates (total district population, primary party, whether an election is an open-seat race, party's voteshare in the district in the 2016 presidential election). We also include specifications that rely on all available automated bandwidth-selection ("bwselection") approaches in the rdrobust() package, as well as varying the order of the polynomial used to construct the point estimates (from 1, local linear, to 5). We include specifications using all available kernel choices from the rdrobust() package. We also vary the types of estimates presented: our main table presents estimates with robust bias-corrected standard errors (Calonico, Cattaneo and Farrell 2020) but in this plot we also include estimates with conventional standard errors. Each set of choices in the plot ("outcomechoice," "subsetchoice," etc.) interacts with all possible values for all other choices, for a total of 1,200 specifications.

The estimates across these specifications are fairly consistent with the paper's main esti-

Figure A5: Exploring many plausible specifications for the main RDD




mate. Of the 1,200 specifications, only a few dozen yield negative point estimates (all non-significant), with the largest negative estimate showing less than three percentage points of disadvantage for minority candidates. The estimates range from -.024 to .25 in size (-2 to 25 percentage points' general-election advantage for minority nominees), and they vary in precision. Most of these estimates rule out even small negative effects: over half of them have 95% confidence intervals that exclude estimates of one percentage point.

It does not appear that bandwidth-selection methods, kernel choice, or bias-correction decisions make much difference for the estimates (these choices are spread across the range of estimates). In general, point estimates look smaller and more precisely-estimated when including covariates, which makes sense given the limited sample size available for this RD analysis.

Inspired by Stommes, Aronow and Sävje (2021), we also consider the statistical power of this design. Our main analysis presents null findings, rejecting claims of an "electoral penalty" for parties that nominate minority candidates. These findings are only compelling to the extent this analysis is powered to detect evidence of such a penalty should it exist. We use the **rdpower()** package to perform ex-ante power calculations for the design presented in the main paper. A local-linear specification including background covariates (as presented in column 2 of the main paper table) has 80% power to detect effects of size .0555 or greater. Such effects (a 5.6-percentage point electoral penalty for nominating a minority candidate) correspond to a cohen's $d$ of .36, a small-to-medium effect size. Being able to rule out effects of this size is substantively meaningful, especially as some experimental work on electoral penalties for nonwhite candidates finds much larger effect sizes. In one recent study, a nationally-representative sample showed about a fifteen-percentage point electoral penalty for nonwhite candidates (in the "face-saving" treatment condition, which the paper suggests is most similar to real-world election conditions) (Krupnikov, Piston and Bauer 2016). In another, respondents in a low-information condition showed huge racial penalties across groups: a 22 percentage point disadvantage for Black candidates and 10 points for Latino or Asian candidates (all compared to white candidates), though these penalties shrank somewhat with the provision of party information (Crowder-Meyer, Gadarian and Trounstine 2020)(Figure 2).

## A4 Relevance of Other Candidate Characteristics

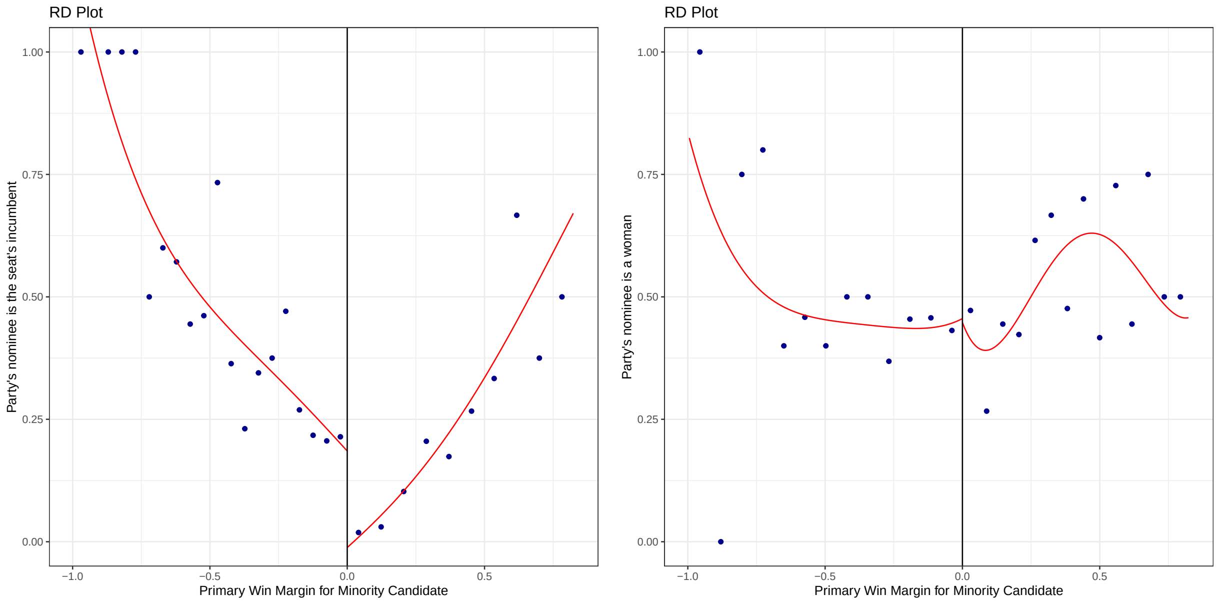
Below we examine how other observable candidate characteristics may also change with candidate race/ethnicity. If we saw, for example, that characteristics like incumbency tended to vary along with winners' race, that could provide an important understanding of the mechanisms by which nominating minority candidates was providing parties with an advantage in the elections we study.

Accordingly, we briefly examine the other candidate characteristics available to us in the dataset. Figure A6 presents two plots similar to our main graphical presentation in Figure 1. In this case, rather than focusing on our main outcome of interest, we instead ask whether nominating a white candidate yields different nominee characteristics on gender or incumbent status. The plot on the left side examines incumbency, and finds a positive discontinuity at the cutpoint: when parties narrowly nominate a white candidate, they are more likely to get a nominee who is also the incumbent for the legislative seat in question. The plot on the right side of the figure examines candidate gender: parties nominating a white candidate for the general election do not appear more or less likely to have nominated a woman.[2] So it appears that when parties narrowly nominate a candidate of color, they are choosing a bundle of characteristics: those narrowly-winning minority candidates are also more likely to be newcomers running for a seat.

The correlation between gender or incumbency and the close nomination of a White candidate reveals potential mechanisms by which candidate race matters for election outcomes. Understanding a person's race to be a "bundled treatment" (Sen and Wasow 2016) from the perspective of estimating causal effects, the fact that barely-nominated white candidates are more likely to be incumbents *who performed worse than a barely-nominated candidate of color* indicates that minority candidates may be more likely to mount a formidable challenge to white incumbents who are strong on paper, but underperform in the general election

[2] Running the RDD specification with "party's nominee is a woman" as the outcome measure yields positive point estimates, suggesting narrowly-winning minority candidates are slightly more likely to be women, but they are statistically indistinguishable from zero.

Figure A6: Examining other candidate characteristics across the cutpoint




campaign. Or, a potential mechanism could be a divisive primary campaign that depresses general election turnout among minority voters.

Another possibility raised by Marshall (2021) is that of "compensating differentials" or of candidates winning the nomination in different ways depending on race. For example, it could be the case that primary voters concerned about candidate electability would strategically discriminate against candidates of color because they feared that these candidates would lose in the general election if nominated (Bateson 2020; Green, Schaffner and Luks N.d.) If this kind of strategic discrimination took place, it could mean that barely-nominated minority candidates would be of higher quality than barely-nominated white candidates (since they had passed through this discriminatory filtering process), and thus could be expected to do better in general elections simply because of the quality differential introduced by our RDD's focus on narrowly-contested elections. Such a pattern would yield statistically-valid estimates, but would only teach us about candidate outcomes in districts with this specific pattern of strategic primary discrimination. If candidate recruitment or selection patterns changed, so too could the general-election vote patterns observed here.

We cannot rule out all possible patterns of compensating differentials because there is limited pre-primary data available on state legislative candidates. However, we do not think such a pattern of "Jackie Robinson" effects is likely in these sorts of low-information elections. For one thing, we note the pattern seen in Figure A6: narrowly-nominated minority candidates are substantially *less* likely to have already been in state legislative office, which is the opposite of what we would expect if there were a quality differential being introduced by the research design.

Second, additional analyses are inconsistent with ideas of strategic discrimination by primary voters. One prominent story about compensating differentials imagines that primary voters are strategic and accurately predict that candidates of color will face bias in the general election. Thus, they hold candidates of color to a higher standard in the primary election, ensuring that any candidates of color that make it to the general election will be more qualified than white nominees, and thus counterbalancing the racial bias of general-election voters. This form of forward-looking strategic discrimination could yield patterns equivalent to the ones shown in our main analyses even if nominees of color did face an electoral penalty in the general. However, we note that this kind of strategic discrimination should only be likely to happen in certain types of elections: specifically, primary voters should be especially likely to engage in strategic discrimination in contexts where the general election is likely to be closely contested. In districts that are "safe" for either party and where a given party's nominee is *ex ante* either very likely or very unlikely to win the general election, there is little reason to engage in strategic discrimination as it is unlikely to make a difference. Thus, we re-run our main RD analysis limiting to primaries in safe partisan districts (defined as those where one party's presidential nominee won more than 60% of the vote in the 2016 election, though a cutoff of 65% or 70% yields equivalent conclusions). Table A4 reproduces the analysis from Table 1 of the main paper within this subset, and finds similar general-election vote patterns (though noisier given the smaller sample). If strategic discrimination were driving the main estimates, we should expect to see negative point estimates in Table

A4: without strategic primary discrimination "masking" the general election bias, we should be able to see any electoral penalty faced by nominees of color. We do not see such a pattern even in the subset of elections where strategic primary discrimination should be least likely.

Table A4: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, Safe-seat races only

| | | | |
|---|---|---|---|
| Minority nominee | 0.16 | 0.19 | 0.37 |
| | (-0.01,0.36) | (0.03,0.41) | (0.02,0.71) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 88 | 84 | 99 |
| Bandwidth | 0.23 | 0.21 | 0.26 |

## A5 Disaggregating Results by Candidate Identity

In the main text we find no evidence that, on average, the close nomination of a minority candidate over a White candidate in a primary election impacts general election partisan outcomes. Our primary design leverages all contests where the first runner-up and winner in state legislative primary elections were of different race/ethnicities and at least one candidate was [non-Hispanic] White. This produces a pool of 400 state legislative primary elections (and their subsequent general elections) from 2018, 2019, and 2020. However, previous research has tended to focus on bias against racial/ethnic minority candidates from one group in isolation, e.g., bias against Black candidates relative to similarly-positioned White candidates.

Table A5: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, Black-White matchups only

| | | | |
|---|---|---|---|
| Minority nominee | 0.17 | 0.01 | 0.02 |
| | (0.06,0.32) | (-0.01,0.04) | (0,0.06) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 90 | 55 | 117 |
| Bandwidth | 0.24 | 0.13 | 0.32 |

Table A6: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, Latino-White matchups only

| | | | |
|---|---|---|---|
| Minority nominee | 0.1 | -0.06 | -0.01 |
| | (-0.12,0.37) | (-0.12,-0.01) | (-0.11,0.1) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 33 | 44 | 47 |
| Bandwidth | 0.13 | 0.19 | 0.23 |

Table A7: RDD Estimates of the Effect of Nominating a Minority Candidate on General-Election Voteshare, Asian-White matchups only

| | | | |
|---|---|---|---|
| Minority nominee | -0.02 | 0.19 | 0.25 |
| | (-0.26,0.28) | (-0.02,0.52) | (-0.09,0.64) |
| Polynomial | p=1 | p=1 | p=3 |
| Covariates | | X | X |
| Effective Sample Size | 14 | 12 | 30 |
| Bandwidth | 0.15 | 0.1 | 0.25 |

In Tables A5 - A7, we conduct our regression discontinuity analysis on subsets of Black-White, Latino-White, and Asian-White primary election matchups in isolation. Across groups and model specifications the confidence intervals in our estimates consistently overlap with the pooled analysis we use in the main text; no one group appears to be consistently responsible for the overall lack of a minority candidate penalty we estimate in the main text. Most point estimates in these tables are positive, indicating that if anything minority candidates outperform white nominees in general elections. Our results for Latino candidates are less consistent, but again overlap with zero effect on average.

That said, we do not take these results to be definitive evidence of the effect for any single group in isolation. As indicated by the effective sample sizes listed in the third rows of Tables A5 - A7, there are simply too few cross-racial primary election matchups over the three election years we examine to make strong claims about differential magnitudes of effects. We look forward to future research that can disaggregate candidate prospects by racial/ethnic group.

# References


Bateson, Regina. 2020. "Strategic discrimination." *Perspectives on Politics* 18(4):1068–1087.

Calonico, Sebastian, Matias D Cattaneo and Max H Farrell. 2020. "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs." *The Econometrics Journal* 23(2):192–210.

Cattaneo, Matias D., Michael Jansson and Xinwei Ma. 2021. *rddensity: Manipulation Testing Based on Density Discontinuity.* R package version 2.2.
**URL:** *https://CRAN.R-project.org/package=rddensity*

Crowder-Meyer, Melody, Shana Kushner Gadarian and Jessica Trounstine. 2020. "Voting can be hard, information helps." *Urban Affairs Review* 56(1):124–153.

Green, Jon, Brian Schaffner and Sam Luks. N.d. "Strategic Discrimination in the 2020 Democratic Primary." . Forthcoming.

Krupnikov, Yanna, Spencer Piston and Nichole Bauer. 2016. "Saving face: Identifying voter responses to black candidates and female candidates." *Political Psychology* 37(2):253–273.

Marshall, John. 2021. "Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics? Working Paper.".

Sen, Maya and Omar Wasow. 2016. "Race as a 'Bundle of Sticks': Designs that Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19.

Stommes, Drew, PM Aronow and Fredrik Sävje. 2021. "On the reliability of published findings using the regression discontinuity design in political science." *arXiv preprint arXiv:2109.14526* .