Can close election regression discontinuity designs identify effects of winning politician characteristics?*

John Marshall†

July 2021

Politician characteristic regression discontinuity (PCRD) designs leveraging close elections are increasingly used to estimate the effects of elected politician characteristics on downstream outcomes. Unlike textbook regression discontinuity designs, treatment is defined by both winning a close election and a predetermined characteristic that could affect their victory margin. I prove that conditioning on the post-treatment victory margin causes PCRD estimators to capture the effect of the specific characteristic of interest and all compensating differentials—candidate-level characteristics that ensure winning candidates remain in close races despite being advantaged/disadvantaged by the characteristic of interest. Isolating this characteristic’s effect generally requires assuming either that the specified characteristic does not affect candidate vote shares or that no compensating differential affects the outcome. Since theories of voting behavior suggest that neither strong assumption usually holds, I further consider whether and how balance tests, covariate adjustment, bounding, and recharacterizing treatment can mitigate the post-treatment bias afflicting PCRD designs.

*I thank Jake Bowers, David Broockman, Naoki Egami, Anthony Fowler, Adam Glynn, Don Green, Andy Hall, Shigey Hirano, Horacio Larreguy, Chris McConnell, Carlo Prato, Pablo Querubín, Alberto Simpser, Tara Slough, Jim Snyder, Jörg Spenkuch, Dan Thompson, Yamil Velez, Yiqing Xu, and attendees at EGAP and Polmeth presentations for helpful feedback and illuminating discussions.

†Department of Political Science, Columbia University. Email: jm4401@columbia.edu.
1 Introduction

Regression discontinuity (RD) designs have become a staple of the quantitative social scientist’s methodological toolkit. RD designs leverage treatment assignments that change discontinuously, once a known threshold in a forcing variable is reached, to identify treatment effects for observations at that threshold. While external validity can be limited, such designs are generally regarded as the observational method best approximating the experimental “gold standard” in terms of internal validity (Lee 2008). As researchers have focused on estimating causal effects, the use of RD designs has exploded over the last decade (de la Cuesta and Imai 2016).

A particularly popular version of the RD design among political scientists and economists uses close elections to estimate effects of a specific elected politician characteristic on downstream electoral, policy, and constituent outcomes. I will call this application a politician characteristic regression discontinuity (PCRD) design. Studies from across the globe have used PCRD designs to compare narrowly-elected politicians that differ in terms of a given predetermined characteristic, usually with the objective of holding observable and unobservable potential confounders constant.¹ The same approach has been used to estimate effects of narrowly-selected primary election winners with different characteristics. Supplementary Table S11 lists 78 published articles, often in prestigious journals, that estimate downstream effects of ascriptive characteristics (gender, race or ethnicity, clan, religious identity), prior actions of politicians (criminal history, prior incumbency, seniority), labels politicians sort into (partisan affiliation, ideology), and institutional status (partisan alignment with other levels of government, term limit status). Indeed, PCRD designs appear to facilitate opportunities to study how electoral selection affects representation, accountability, and participation that are only limited by a researcher’s capacity to measure politician characteristics of interest.²

While the appeal of credibly estimating effects of winning candidate characteristics is obvious,

---

¹The target estimand is rarely stated explicitly, although most studies imply that PCRD designs identify an “all else equal” effect.
²Since multiple testing concerns could be addressed using standard tools, this is not this article’s primary focus.
whether PCRD designs can isolate the effect of a specific $X$—the characteristic, or bundle of characteristics, of interest—around the discontinuity is not. In this article, I demonstrate when and how a non-standard feature of this application of the RD design causes PCRD designs to generally identify compound treatment effects, rather than the local average treatment effect (LATE) of $X$. This confounding reflects a form of post-treatment bias, and is avoided by imposing one of two strong—and often implausible or unverifiable—assumptions. I further consider the merits and limitations of strategies to combat or reconceptualize this potential bias.

The source of bias lies in the difference between standard RD and PCRD designs. Standard RD designs define treatment as falling above or below a threshold. For example, close elections have been used to vary whether a candidate or party was elected or not to identify financial returns to holding office (Eggers and Hainmueller 2009) and incumbent party electoral advantages (Lee 2008). In contrast, the treatment in PCRD designs—which instead seek to estimate the LATE of an elected politician characteristic—is defined by possessing (or not) predetermined characteristic $X$, conditional on narrowly winning an election. Beyond targeting different estimands, the mechanics of PCRD designs differ from standard RD designs in two important ways. First, and most obviously, $X$ may be unconditionally correlated with other candidate-level characteristics. Such correlations between characteristics create the risk of confounding, or necessitate reinterpreting the estimand as a compound treatment. Second, and more subtly, restricting attention to close elections entails conditioning on candidate vote shares that may be affected by $X$. As Table 1 below demonstrates, the former difference is increasingly acknowledged by researchers, but the latter is not. This article focuses on the second issue, showing that PCRD designs induce bias, even when $X$ is unconditionally uncorrelated with other predetermined variables and the weak continuity assumption underpinning standard RD designs holds.

By expressing PCRD designs in terms of standard RD designs, I show that conditioning the sample on close elections causes PCRDs to identify the LATE of electing a candidate of type $X$ combined with a (differential-weighted) LATE of any compensating differentials. While PCRD designs ensure balance across districts different types of candidates are elected from, compensating
differentials are characteristics of individual candidates that: (a) the researcher defines as distinct from $X$; and (b) cause candidates of type $X$ to still be in close elections. For example, in seeking to isolate the effect of gender, competence could be a compensating differential if women in close elections are more competent than men in close races because voters are biased against women. While determining which correlated characteristics should be regarded as part of the treatment is not always obvious, I abstract from this study-specific decision by assuming that there exist some compensating differentials $Z$ that are regarded as distinct from the bundle of characteristics that define $X$.

I further establish sufficient conditions that isolate the effect solely attributable to characteristic $X$ and characterize the bias when these strong additional assumptions fail to hold. Specifically, the PCRD estimator yields an unbiased estimate of $X$’s effect if—at the discontinuity—either: (i) $X$ does not affect the winning candidate’s victory margin (the forcing variable); or (ii) none of the compensating differentials in the vector $Z$ affects the outcome of interest $Y$. These conditions are analogous to the condition under which the bias associated with conditioning the sample on a post-treatment variable disappears (e.g. Hernán and Robins 2011; Montgomery, Nyhan and Torres 2018). Moreover, where neither condition holds, I show that compensating differentials cause PCRD designs to underestimate the LATE of $X$ when $X$ and the net impact of $Z$ affect both the candidate’s vote share and the outcome $Y$ in the same direction.

I highlight three practical implications for applied research. First, researchers should explicitly state and support one of the additional conditions in order to claim that PCRD estimates can isolate the effect of $X$ by design. However, these assumptions are often difficult to empirically substantiate and are theoretically implausible where voters observe $X$ and believe it will affect outcomes they care about. Second, where neither condition can be sustained, strategies for mitigating threats to internal validity vary in their effectiveness. My analytic results show how non-null PCRD estimates can be combined with (dis)continuity tests to bound the LATE of $X$. Conversely, continuity test and covariate adjustment strategies struggle to address the biases that PCRD induce because covariate discontinuities should be expected in PCRD designs, even after covariate adjustment.
Third, although isolating the effect of $X$ is often desirable from a theoretical or policy perspective, researchers might consider reinterpreting PCRD estimates as a compound treatment incorporating the effects of all compensating differentials. Where researchers can measure plausible compensating differentials, (dis)continuity tests can facilitate interpretation of this expanded, albeit less well-defined and possibly heterogeneous, conception of treatment.

In clarifying the interpretation of PCRD designs seeking to isolate effects of winning candidate characteristics, this article makes several contributions. First, I provide the first systematic account of the challenges that arise when treatment in RD designs is defined by a variable that may also affect the forcing variable. While some articles have noted study-specific issues relating to how conditioning on close elections introduces compensating differentials (e.g. Gagliarducci and Pascerman 2012), 71 of the 78 of the published studies using PCRD designs—as well as a recent review of close election RD designs (de la Cuesta and Imai 2016)—do not mention the possibility of compensating differentials. Second, I move beyond demonstrating the challenges in interpreting PCRD designs by deriving additional assumptions under which the LATE of $X$ can be isolated, establishing when PCRD designs underestimate and overestimate the LATE of $X$, and proposing strategies that facilitate the interpretation of PCRD results. Third, despite the agnostic approach underpinning my statistical analysis, this article also reiterates the important role for theory in interpreting the estimands identified by a popular research design. By highlighting the need to understand why candidates end up in close elections, this article complements a growing literature emphasizing the theoretical implications of empirical models (e.g. Bueno de Mesquita and Tyson 2020; Eggers 2017).

The issues that I highlight are distinct from prior critiques of RD designs leveraging close elections. Extant studies have examined other ways through which compound treatments can confound causal attribution across variables, including where multiple treatments are assigned at the same threshold (Eggers et al. 2018), particular characteristics—like Black politicians in the U.S. overwhelmingly being Democrats—being bundled together (e.g. Bucbianeri 2018; Ferreira and
Gyourko 2014; Hall 2015; Hopkins and McCabe 2012), and where treatment affects downstream behaviors such as future candidacy decisions (e.g. Eggers 2017). I instead emphasize how the definition of treatment variables more fundamentally creates compound treatments from predetermined variables by conditioning on close elections. Others have debated whether election outcomes are truly determined by chance at the discontinuity (see Caughey and Sekhon 2011; de la Cuesta and Imai 2016; Eggers et al. 2015) and highlighted the sensitivity of RD estimates and inference to functional form assumptions (Gelman and Imbens 2019). However, the conceptual problems raised by this article still arise when the standard RD assumption of no sorting at the threshold holds and enough data exists for non-parametric estimation at the threshold.

2 The use of PCRD designs

This section builds intuition for the issues that arise when PCRD designs are used to estimate effects of elected politician characteristics. I first describe the design and potential problems through the lens of two prevalent examples—estimating effects of gender and party affiliation. I then review published articles to illustrate the prevalence of PCRD designs and summarize how these designs have been implemented.

2.1 Examples of PCRD designs

2.1.1 Electing women

I start with the example where researchers compare differences in outcomes between polities where women are elected relative to men. Recent studies have used PCRD designs to estimate effects of electing women on policy priorities (e.g. Clots-Figueras 2011; Ferreira and Gyourko 2014), corruption and hiring practices in office (Brollo and Troiano 2016), turnout among women and women running for office at future elections (e.g. Broockman 2014; Ferreira and Gyourko 2014), and gov-

3While my analysis also highlights how groups of characteristics come together, the form of post-treatment bias that I identify still applies once a treatment that differentiates bundles of related characteristics has been delineated.
ernment instability (Gagliarducci and Paserman 2012). Others have similarly examined the effect of women winning primary elections on financial contributions (Anastasopoulos 2016) and general election results (Buchianeri 2018). Most of these studies describe their estimand as identifying the effect of electing a woman instead of a man, often implicitly holding other characteristics of the candidate constant. In the more explicit formulations of this claim, (Brollo and Troiano, 2016:28) describe their study as estimating “the causal effect of the gender of the policymaker,” while (Ferreira and Gyourko, 2014:24) describe their design as estimating the “effect of gender” and (Clots-Figueras, 2011:665) describes her design as identifying “the effect of a legislator’s gender.”

Before illustrating the empirical issue at hand, it should be emphasized that defining gender as a treatment that is conceptually distinct from other candidate characteristics is challenging. This is because gender is often viewed as an inherently bundled treatment comprising various correlated features (Sen and Wasow 2016); women that win close elections may espouse different policies, possess different qualifications, or have different personalities from men that win close elections. To isolate the effect of electing a women from potential confounds, the researcher must distinguish the bundle of characteristics that conceptually differentiate women and men candidates—the definition of treatment—from the characteristics they regard as distinct from the treatment—the potential confounders. Gender is a particularly challenging example, but the need to explicitly define treatment applies equally to other characteristics—such as prior experience or partisan alignment with other relevant politicians—that may be easier to separate conceptually.

Once a researcher has identified a set of characteristics that are conceptually distinct from a candidate’s gender, PCRD designs then estimate the politician gender in single-member plurality races by comparing outcomes in “treated” districts where a woman was just elected in a race against a man with outcomes in “control” districts where a man was just elected in a race against a woman. By focusing on close elections, it is usually argued that the two types of districts—with respect to both district-level characteristics and the individual-level characteristics of winning candidates—will be identical, in expectation, in all other ways. If this were the case, the design would identify

---

4Gender is also hard—if not impossible—to manipulate. I use this example because of its prevalence in the literature.
the effect of electing a woman over electing an otherwise-similar man. To ease exposition, I will call the bundle of other—not gender-related—characteristics “competence” and assume that voters prefer more competent candidates because they achieve positive district outcomes.

Figure 1a plots hypothetical data for the cases that researchers observe, as a function of victory margin, in black. The cases to the right of the vertical line, where a woman is elected ahead of a man, are treated. The relevant counterfactual, shown in gray, is equally competent men who win other elections against women. The difference between elected men and women at the discontinuity, $\tau$, is the true LATE of having elected a woman over a man in a close election.

Is it reasonable to assume at the candidate level that men that narrowly win are equally competent as women that narrowly win? Suppose that—holding competence fixed—voters are more likely to vote for men (Anzia and Berry 2011; Lawless 2015); this could arise from stereotyping, varied media attention, or differential support from political elites.\(^5\) Conditional on a man being in a close race with a woman, the man must then possess lower levels of competence in expectation than the woman against whom they competed. The observed outcomes among men that won are

---

\(^5\)I make this assumption for illustrative purposes; recent evidence suggests that voter biases may be subsiding (Dolan 2014; Hayes and Lawless 2016).
shown in black to the left of the vertical line. The PCRD estimate $\hat{\tau}$ is then an upwardly biased estimate of $\tau$ in this example because it is partially confounded by competence—the compensating differential required for women to be in close races with men when voters are biased against women.

2.1.2 Party affiliation

Analogous challenges apply to estimating effects of a candidate’s party affiliation. Studies using PCRD designs compare outcomes between political units that elected candidates from different parties (e.g. Gerber and Hopkins 2011; Lee, Moretti and Butler 2004; Pettersson-Lidbom 2008). These studies describe the design as capturing the “effect of a Democratic victory” (Gerber and Hopkins 2011:335) or “causal estimates of the effect of party control” (Pettersson-Lidbom 2008:1037). This example focuses on electing Democrats over Republicans to office.

The researcher must again specify what does and does not constitute part of the partisan affiliation treatment. To illustrate, I assume that party affiliation captures a common set of policy positions or ideology and that Democrats are more popular with voters on average. While researchers might reasonably conceive of selection into parties as part of their treatment, I again consider candidate “competence” as the potential compensating differential. Other compensating differentials could include alignment with higher-level incumbents or prior performance in office.

As with the gender example, PCRD designs compare districts where Democrats and Republicans barely won close elections against a candidate from the other party. Since Democrat candidates are more popular than Republican candidates, Republican candidates need to be relatively more competent to counteract this disadvantage. By conditioning on close elections, the PCRD design then compares a Democrat that is atypically incompetent with a Republican that is atypically competent. As Figure 1b illustrates, the PCRD estimate $\hat{\tau}$ is now downwardly biased because narrowly-elected Democrats are less competent than narrowly-elected Republicans.


2.2 Limited awareness of compensating differentials

To examine awareness of these potential issues, I identified 78 published articles using PCRD designs.\textsuperscript{6} The earliest article was published in 2004; reflecting the design’s growing popularity, 71\% of the articles have been published since 2015. These articles have consistently appeared in prominent journals in political science and economics: 24\% were published in the \textit{American Journal of Political Science}, \textit{American Political Science Review}, or \textit{Journal of Politics}, while a further 9\% were published in the \textit{American Economic Review}, \textit{Econometrica}, the \textit{Quarterly Journal of Economics}, or the \textit{Review of Economic Studies}. According to Google Scholar, these studies had collectively amassed 9,832 citations by July 21, 2021.

After reading each article, I hand-coded whether the article demonstrated awareness of four potential threats to internal validity. Specifically, I coded whether an article: (i) assessed continuity in potential outcomes by testing for discontinuities in other covariates; (ii) assessed the same continuity assumption using density tests for sorting; (iii) recognized that candidate characteristics may come as bundles due to unconditional correlations between characteristics; and (iv) discussed the risk of inducing or altering correlations between candidate characteristics by conditioning on close elections. The first two threats are benchmarks that apply to all RD designs, the third could apply to other RD designs but is particularly relevant for PCRD designs, and the fourth issue is specific to PCRD designs. My coding of awareness was generous, such that brief and suggestive references to an issue—sometimes just in footnotes—were coded positively.

The results in Table 1 indicate that applied researchers are generally aware of the importance of validating the continuity assumption and, to a lesser degree, that candidate characteristics come as bundles that are hard to separate. Panel A shows that 90\% of articles conducted balance or continuity tests and 68\% explored potential sorting around the threshold using density tests like the one proposed by McCrary (2008). Panel B shows that both strategies for validating the continuity assumption have become more prevalent over time.\textsuperscript{7} Only 36\% of articles demonstrated awareness

\textsuperscript{6}I exclude the many unpublished articles using PCRD designs.

\textsuperscript{7}This claim is based on a bivariate regression of demonstrating awareness of an issue on the year of publication.
Table 1: Number of studies using PCRD designs that demonstrate awareness of different threats to internal validity

<table>
<thead>
<tr>
<th></th>
<th>Number of articles</th>
<th>Covariate continuity tests</th>
<th>Sorting/density tests</th>
<th>Bundled characteristics</th>
<th>Compensating differentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: All articles</td>
<td>78</td>
<td>70</td>
<td>53</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Panel B: Articles by five-year period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002-2006</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2007-2011</td>
<td>9</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2012-2016</td>
<td>33</td>
<td>28</td>
<td>23</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>2017-forthcoming</td>
<td>35</td>
<td>35</td>
<td>29</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Panel C: Articles by topic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criminal history</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Education</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Gender</td>
<td>14</td>
<td>13</td>
<td>9</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Ideology</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Incumbency, term limit status, or seniority</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Partisan alignment across government tiers</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Partisan affiliation</td>
<td>33</td>
<td>30</td>
<td>21</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Pre-office vocation</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Race, ethnicity, religion, or clan</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Panel D: Articles by field</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economics</td>
<td>42</td>
<td>38</td>
<td>32</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>Political Science</td>
<td>35</td>
<td>31</td>
<td>20</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Sociology</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: This list, which excludes unpublished studies and may be incomplete, should be considered a lower bound on this design’s usage.

of the possibility that the candidate characteristic of interest might be correlated with other characteristics. Surprisingly few studies recognized that types of candidates may differ across political parties. However, as panel C demonstrates, the majority of studies seeking to estimate effects of candidate gender, incumbency status, vocation, education, and ideology at least briefly discussed this issue.

In contrast, very few studies demonstrated any awareness of the issue motivating this article—the risk of inducing or altering correlations between candidate characteristics by restricting the sample to close elections. Indeed, only 7 of the 78 articles even loosely mention this issue. Most
of these articles briefly note the possibility that elections may be close between two candidates that
differ in terms of the characteristic of interest because they could be counterbalanced by differences
in terms of another characteristic; in 5 of 7 cases, this arises from the specific concern—that is
well-recognized by scholars studying women in politics (Lawless 2015)—that women in close
races may differ from men due to voter biases or campaign disadvantages. However, the origin
and implications of such compensating differentials received limited discussion and were often
quickly dismissed, despite the fact—as I will soon show—that compensating differentials between
candidates are generally likely to be required in PCRD designs.

Only one article delves deeper into the issue. (Gagliarducci and Paserman, 2012:1031) note
that identifying the effect of elected women using a PCRD design requires that “the vote share of
each candidate must not depend directly on gender.” They suggest that this assumption is plausi-
ble in their study examining the effect of women narrowly elected as mayors on early municipal
government termination in Italy, if voters do not select candidates to maintain government stabil-
ity, are unaware that mayor gender affects government stability, or only select candidates on the
basis of factors unrelated to gender that could affect government stability. This article generalizes
the conditions—which include, but are not limited to, those identified by Gagliarducci and Paser-
man (2012)—under which the PCRD estimator yields unbiased estimates of the desired estimand,
characterizes the nature of bias when these strong conditions do not hold, and discusses mitigation
strategies and alternative conceptualizations of treatment.

3 Theoretical analysis

This section first recaps how standard RD designs work in the context of close plurality elections
in single member districts, before explaining how PCRD designs differ. I then show that the
differences between the designs can bias PCRD estimates of the effect of the characteristic of in-

---

8RD designs have been adapted to leverage discontinuities in the share of close races won by a particular type of
candidate within a district (Clots-Figueras 2011) and, in the context of proportional representation systems, the number
of seats received (Folke 2014) and majority control of legislative councils (Pettersson-Lidbom 2008) by particular
parties. While these designs differ in some respects, the same challenges enumerated below apply when used to isolate
effects of elected politician characteristics.
interest in a stylized example. I finally provide general theoretical results demonstrating that this bias will emerge unless additional assumptions—that are far stronger than the standard RD continuity assumption—are imposed.

3.1 Standard RD designs

In the close election application of RD designs, each candidate $i$ in district $d$ receives vote share $V_{id} \in [0,1]$. The forcing variable is the difference between $i$’s vote share and the vote share $V_{jd}$ of the most popular other candidate $j \neq i$ in the district: $\Delta_{id} := V_{id} - V_{jd} \in [-1,1]$. The following binary treatment variable then indicates whether candidate $i$ won the election in district $d$ or not:

$$T_{id} := \begin{cases} 1 & \text{if } \Delta_{id} > 0, \\ 0 & \text{if } \Delta_{id} \leq 0. \end{cases}$$ (1)

In addition to observing $T_{id}$ based on which candidate wins the race, researchers also observe an outcome variable $Y_{id}$ for each candidate. The potential outcome $Y_{id}(T_{id}) \in \mathbb{R}$ depends on the candidate’s treatment assignment, and has a finite expectation. This representation encodes the SUTVA assumption that $i$’s potential outcome is not affected by the treatment status of other candidates and that there is a single version of treatment. Since only one potential outcome can be observed, the observed outcome is related to potential outcomes by $Y_{id} = T_{id}Y_{id}(1) + (1 - T_{id})Y_{id}(0)$.

The standard RD design requires the following weak continuity assumption:

**Assumption 1.** Potential outcomes $Y_{id}(T_{id})$ satisfy the following conditions:

(a) *Continuity from above:* $\lim_{v \downarrow 0} \mathbb{E}[Y_{id}(1) | \Delta_{id} = v] = \mathbb{E}[Y_{id}(1) | \Delta_{id} = 0]$;

(b) *Continuity from below:* $\lim_{v \uparrow 0} \mathbb{E}[Y_{id}(0) | \Delta_{id} = v] = \mathbb{E}[Y_{id}(0) | \Delta_{id} = 0]$.

The assumption states that, at the point of discontinuity, potential outcomes do not vary discontinuously in any way other than whether a given candidate won the election. This implies that all

9I assume that $i$ does not win if $V_{id} = V_{jd}$. All results hold if $i$ and $j$ each win with probability 0.5 when $V_{ij} = V_{jd}$. 

13
characteristics of candidates or districts other than $T_{id}$ vary continuously with $\Delta_{id}$ as a candidate passes the threshold from loser to winner. In the case of close elections, this is plausible because idiosyncratic factors exogenous to candidate characteristics, such as election day weather, can generate random variation in which candidate wins; Eggers et al. (2015) provide evidence supporting this claim from ten countries across the world.

Where continuity holds, it is natural to estimate the LATE of $T_{id}$ at the point of discontinuity—denoted formally by $\tau_{RD} := \mathbb{E}[Y_{id}(1) - Y_{id}(0) | \Delta_{id} = 0]$—by comparing observed outcomes between candidates that narrowly won and narrowly lost. The standard RD estimator is given by:

$$\hat{\tau}_{RD} = \lim_{v \downarrow 0} \mathbb{E}[Y_{id} | \Delta_{id} = v] - \lim_{v \uparrow 0} \mathbb{E}[Y_{id} | \Delta_{id} = v],$$

where the limits of each conditional expectation at the point of discontinuity are usually estimated by means within small bandwidths, a local linear regression within small bandwidths, or flexible functions within larger bandwidths. Throughout, I abstract from the estimation problem by assuming that unbiased estimates for each limit have been obtained.\(^\text{10}\) Consequently, standard results (e.g. Imbens and Lemieux 2008) show that the RD estimator is unbiased:

**Proposition 1.** $\hat{\tau}_{RD}$ is an unbiased estimator of the LATE of $T_{id}$ at $\Delta_{id} = 0$.

**Proof:** see Supplementary Information for all proofs. ■

This type of RD design has proved popular for estimating the consequences of being elected. One strand of this literature has explored the effect of being elected to office on downstream wealth (e.g. Eggers and Hainmueller 2009). Another strand has studied the effect of winning elections on subsequent election outcomes by examining whether winners—or those linked to them—run in and win future elections (e.g. Querubín 2016). Since the decision to run in future elections often depends on winning a close election, researchers interested in isolating the effect of winning office, conditional on running, on future electoral success have focused on the electoral outcomes of parties linked to the winning and losing candidates where parties always run for office. This

\(^{10}\)See Cattaneo, Idrobo and Titunik (2019) for a recent discussion of estimation strategies.
has spawned a large literature measuring party incumbency advantages.\textsuperscript{11} In the U.S., many such studies find that Democrats are more likely to win future elections in districts narrowly won by a Democrat (e.g. Lee 2008). These studies are united by comparing downstream outcomes between election winners and losers.

3.2 How PCRD designs differ from standard RD designs

Whereas standard RD designs applied to close elections estimate effects of being elected, PCRD designs instead seek to isolate the effect of a specific characteristic of elected politicians on downstream outcomes. Rather than compare winning and losing candidates, PCRD designs compare politicians who all narrowly won elections in different districts but differ according to a predetermined binary characteristic—or bundle of characteristics—denoted by $X_{id} \in \{0, 1\}$.\textsuperscript{12} Characteristics of empirical interest have included gender, race, vocational experience, criminality, prior incumbency, partisan affiliation, and partisan alignment.

Isolating the effect of a characteristic of interest is difficult because politicians are defined by many characteristics that tend to be correlated. For example, elected women in the U.S. are more likely to be Democrats, politicians that have engaged in corruption are more likely to be aligned with higher-level politicians, or politicians from traditional parties are more likely to be experienced. If the characteristic—or bundle of characteristics—of interest $X_{id}$ is correlated with a vector of $K$ other candidate-level characteristics $Z_{id} \in \mathbb{R}^K$, any effect of $X_{id}$ may be confounded by the effects of $Z_{id}$. To define their target estimand, researchers must decide which characteristics should and should not be included in their treatment of interest; put differently, they must decide which characteristics are conceptually distinct from the characteristic of interest. This is an inexact science. Some researchers (explicitly or implicitly) assert a characteristic of interest and seek to substantiate the claim that only this characteristic drives any effect. Others define their treatment

\textsuperscript{11}Incumbency advantage studies using RD design usually focus on parties, rather than individuals, because individual winners and losers rarely both compete at the next election.

\textsuperscript{12}The characteristic is defined in discrete terms to generate distinct treatment conditions. Non-binary characteristics could compare any two characteristic values or create bins.
as a bundle of correlated characteristics that are distinct from some other characteristics.\textsuperscript{13}

The challenge of interpreting compound treatments is recognized in some of the extant literature (see Table 1). To abstract from this issue, I will at times assume that $X_{id}$ is independent of $i$’s conceptually-distinct characteristics $Z_{id}$ and the conceptually-distinct characteristics $Z_{jd}$ of their chief competitor $j$ among politicians in close elections:

**Assumption 2.** At the point of discontinuity, $X_{id}$ is independent of $Z_{id}$ and $Z_{jd}$.

Under this assumption, I show that biases emerge in PCRD designs even in a “best case scenario” where $X_{id}$ is unconditionally independent of $Z_{id}$ and $Z_{jd}$ at the point of discontinuity. This assumption is relaxed for the most general theoretical results.

In shifting attention to the type of politician that wins, the unit of analysis in PCRD designs is the district. The district-level forcing variable then $\Delta_d := V_{1d} - V_{0d} \in [-1,1]$, where $V_{1d}$ and $V_{0d}$ respectively denote the vote shares of the most popular politician of type $X_{id} = 1$ and $X_{id} = 0$. This difference in vote shares is not defined for elections in districts where the top two candidates were of the same type. The corresponding treatment indicates whether a candidate of type $X_{id} = 1$ won the election:

$$X_d := \begin{cases} 
1 & \text{if } \Delta_d > 0, \\
0 & \text{if } \Delta_d \leq 0. 
\end{cases} \quad (3)$$

District-level potential outcomes are similarly defined by the type of politician that won: $Y_d(X_d) = X_d Y_{1d}(1) + (1 - X_d) Y_{0d}(1)$. For example, $Y_d(1)$ and $Y_d(0)$ could correspond to the district-level outcome if the elected candidate was a woman and a man, respectively. Let $Y_{id}(1) = y(X_{id}, Z_{id}) + \nu_d$, where $y$ is a general function and district-level noise $\nu_d \in \mathbb{R}$ is drawn independently of all other variables. In PCRD designs comparing observations of $Y_d$ across districts, the potential outcome for a politician that loses the election, $Y_{id}(0)$, is neither relevant nor well-defined because the loser cannot affect district-level outcomes because they do not enter office.

\textsuperscript{13}In extreme cases, some researchers view all characteristics as a single bundle; I return to this interpretation, to which the following analyses do not apply, below.
The LATE of interest in PCRD designs is the difference in potential outcomes across districts with close elections where politicians of different types were elected: \( \tau_{PCRD} := \mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0] \). This is typically estimated using the following estimator:

\[
\hat{\tau}_{PCRD} = \lim_{v \downarrow 0} \mathbb{E}[Y_d|\Delta_d = v] - \lim_{v \uparrow 0} \mathbb{E}[Y_d|\Delta_d = v]
\]

\[
= \lim_{v \downarrow 0} \mathbb{E}[Y_{id}(1)|\Delta_{id} = v, X_{id} = 1, X_{jd} = 0] - \lim_{v \uparrow 0} \mathbb{E}[Y_{id}(1)|\Delta_{id} = v, X_{id} = 0, X_{jd} = 1].
\]

I again assume that unbiased estimates exist for each conditional expectation. The second line rewrites the district-level estimator in terms of candidate-level potential outcomes to distinguish the standard RD and PCRD designs: whereas standard RD designs compare candidates either side of the threshold in the forcing variable, PCRD designs compare winning candidates that differ in terms of \( X_{id} \) and thus condition on a predetermined difference between \( X_{id} \) and \( X_{jd} \) that could also affect the forcing variable \( \Delta_d \) in addition to \( \Delta_{id} = 0 \). I next show how this distinction creates the opportunity for bias.

### 3.3 The bias of PCRD designs with a single compensating differential

To build intuition, I start with the simple case where a single compensating differential \( Z_{id} - Z_{jd} \) ensures that the race between candidates \( i \) and \( j \) in district \( d \) is close despite only one candidate possessing characteristic \( X_{id} \). Without loss of generality, assume that characteristic \( X_{id} \) helps a candidate win votes, e.g. by being the incumbent, representing a popular party, or by not suffering gender-based biases. The compensating differential offsets this advantage, e.g. because the candidate for whom \( X_{id} = 1 \) is less competent, more malfeasant, or less politically connected than the candidate for whom \( X_{id} = 0 \). By Assumption 2, \( X_{id} \) and \( (Z_{id}, Z_{jd}) \) are unconditionally independent at the point of discontinuity. Consequently, any bias in the PCRD estimator must emerge from conditioning on close races between candidates of different types.

To serve as a compensating differential, \( Z_{id} - Z_{jd} \) must affect the candidate’s vote share and may, if elected, also affect the district-level outcome. I capture this through the following functional
forms:

\[ V_{id} = \alpha \frac{X_{id} - X_{jd}}{2} + \beta \frac{Z_{id} - Z_{jd}}{2} + \frac{\epsilon_{id} - \epsilon_{jd}}{2}, \quad (5) \]

\[ y(X_{id}, Z_{id}) = \tau X_{id} + \gamma Z_{id}, \quad (6) \]

where \( \alpha \geq 0 \) and \( \beta > 0 \) imply that possessing more of characteristic \( X_{id} \) or \( Z_{id} \) increases a candidate’s vote share,\(^{14}\) while \( \tau \) and \( \gamma \) respectively denote the (constant) effects of \( X_{id} \) and \( Z_{id} \) on district-level outcomes. I further assume that \( Z_{id} - Z_{jd} \sim N(0, \sigma_Z^2) \) and \( \epsilon_{id} - \epsilon_{jd} \sim N(0, \sigma_\epsilon^2) \) are normally distributed and drawn independently of both \( X_{id} \) and each other; these distributional assumptions embed Assumption 2.

### 3.3.1 Derivation of bias

Although \( X_{id} \) is unconditionally uncorrelated with compensating differentials, the PCRD design can generate such a correlation by conditioning on close elections where two narrow winners with different characteristics obtain similar vote shares. To see why, first note that the limiting case of close elections—where candidates within a given district receive equal numbers of votes—implies:

\[ V_{1d} = V_{0d} \iff \alpha + \beta (Z_{1d} - Z_{0d}) + \epsilon_{1d} - \epsilon_{0d} = 0, \quad (7) \]

where the candidate of type \( X_{id} = 1 \) is denoted by \( i = 1 \) and the candidate of type \( X_{id} = 0 \) is denoted by \( i = 0 \). The occurrence of any close election between these candidate types implies that candidate 1 ties their race because there is a countervailing compensating differential \( (Z_{1d} < Z_{0d}) \) and/or because they encountered unfortunate random shocks \( (\epsilon_{1d} < \epsilon_{0d}) \).

Conditional on such close races, the PCRD estimator then identifies the following quantity in

\(^{14}\)Both coefficients are positive for simplicity, but the same logic applies for unrestricted \( \alpha \) and \( \beta \). It is natural to consider these coefficients as common across candidates because differences across candidates are characterized by \( X_{id} \) and \( Z_{id} \).
expectation:

\[
\mathbb{E}[\hat{\tau}_{\text{PCRD}}] = \mathbb{E}[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] \\
= \tau + \gamma \mathbb{E}[Z_{id} - Z_{0d} | \alpha + \beta (Z_{1d} - Z_{0d}) + \epsilon_{1d} - \epsilon_{0d} = 0] \\
+ \mathbb{E}[\nu_d | \Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[\nu_d | \Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] \\
= \tau - \gamma \frac{\alpha \beta \sigma_Z^2}{\sigma_\epsilon^2} \left( 1 + \frac{\beta^2 \sigma_Z^2}{\sigma_\epsilon^2} \right),
\]

(8)

where the first line follows from Assumption 1(a) and the district-level outcome being the potential outcome associated with the winning candidate in that district, the second line substitutes in the functional forms for potential outcomes and vote shares, and the final line applies the distributional assumptions on \(Z_{id} - Z_{jd}, \epsilon_{id} - \epsilon_{jd},\) and \(\nu_d.\)

Since PCRD designs compare candidates that win elections, only part (a) of Assumption 1 is needed. This technically weakens the continuity assumption, although it is difficult to imagine contexts where part (a) held but part (b) did not.

This derivation shows how a form of post-treatment bias can cause PCRD designs to yield biased estimates of the effect of characteristic \(X_{id}.\) Where \(X_{id}\) affects \(V_{id},\) the estimator \(\hat{\tau}_{\text{PCRD}}\) conditions on an event \(\Delta_{id} = 0\) that may itself be caused by the treatment. Satisfying \(\Delta_{id} = 0\) when \(X_{id} \neq X_{jd}\) generally requires that compensating differentials—other features of candidates that may in turn influence outcomes—also differ in expectation. Similar insights emerge where there are multiple compensating differentials, as section SI.3 illustrates in the case of two correlated compensating differentials.

Bias can be avoided if one of the following three conditions holds. First, the PCRD estimator is unbiased when no compensating differential is needed \((\alpha = 0).\) This requires voters not to select politicians on the basis of \(X_{id}.\) Second, there is no bias when the compensating differential does not

---

15 The first line also applies Lebesgue’s Convergence Theorem. Since \(Z_{id} - Z_{jd}\) and \(\epsilon_{id} - \epsilon_{jd}\) are normally distributed, \(\mathbb{E}[Z_{idk} - Z_{jd|\Delta_d} = \mathbb{E}[Z_{idk} - Z_{jd|\Delta_d}] + \frac{\text{Cov}[Z_{idk} - Z_{jd|\Delta_d} - \mathbb{E}[\Delta_d]]}{\text{Var}[\Delta_d]}(\Delta_d - \mathbb{E}[\Delta_d])\) must hold. In this application, \(\mathbb{E}[Z_{idk} - Z_{jd}] = 0, \text{Cov}[Z_{idk} - Z_{jd|\Delta_d}] = \beta^2 \sigma_Z^2, \text{Var}[\Delta_d] = \beta^2 \sigma_Z^2 + \sigma_\epsilon^2,\) and \(\Delta_d - \mathbb{E}[\Delta_d] = -\alpha.\) By independence, \(\mathbb{E}[\nu_d | \Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[\nu_d | \Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] = 0.\)
affect the outcome ($\gamma = 0$). This occurs when $Z_{id} - Z_{jd}$ affects which candidate voters prefer but does not ultimately affect the district outcome of interest. A third scenario where bias is avoided occurs when close elections only arise where the difference in vote share due to $X_{id}$ is exactly offset by noise that does not affect district outcomes ($\alpha + \epsilon_{1d} - \epsilon_{0d} = 0$). This condition is closely related to the previous condition, since noise is defined by factors external to a candidate that affect vote choices without affecting district outcomes, and is only plausible where the signal to noise ratio ($\frac{\sigma_Y^2}{\sigma^2}$) is sufficiently low that candidate vote shares are predominantly determined by chance, rather than systematic differences in expected attributes. As I discuss below, theories of voting behavior suggest that these conditions are rarely likely to hold.

### 3.3.2 How does the direction and magnitude of bias vary?

The bias of PCRD estimates can be upward or downward, depending on how $X_{id}$ and $Z_{id}$ affect district outcomes. Where the effects of each variable agree, that is to say either $\tau, \gamma > 0$ or $\tau, \gamma < 0$, equation (8) shows that the PCRD estimator is downwardly biased in magnitude. A PCRD estimate will thus underestimate the effect of $X_{id}$ when both candidate characteristics appeal to voters and also both lead to better (or worse) district-level outcomes once a politician enters office. This would occur where voters select candidates with characteristics that they correctly anticipate will produce better outcomes in office from the perspective of most voters, such as greater economic performance, security, or redistribution towards the majority group. Intuitively, the PCRD estimate is then a lower bound on the effect of $X_{id}$ because differences in $Y_d$ due to electing a candidate possessing desirable characteristic $X_{id}$ are counteracted by electing a candidate possessing relatively less of the also-desirable compensating differential $Z_{id}$. The partisan affiliation example above illustrates such a case.

PCRD designs will instead overestimate the effect of $X_{id}$ where the sign of the two effects disagree. As illustrated by the gender example above, opposing effects could arise when voters misconstrue one of these characteristics to be desirable when it actually harms the outcomes that voters care about, e.g. if voters incorrectly believe that women will perform worse. Opposing effects
could similarly occur where vote buying efforts win votes for candidates whose non-programmatic policies later reduce voter welfare.

Each form of bias creates different challenges for interpreting hypothesis tests. Where underestimation occurs, the direction of a PCRD estimate that rejects the null hypothesis remains correct because \( \hat{\tau}_{PCRD} \) is a lower bound. Conversely, a failure to reject the null hypothesis is relatively uninformative because \( \hat{\tau}_{PCRD} \approx 0 \) is consistent with both a null effect of \( X_{id} \) and a positive effect of possessing \( X_{id} \) canceling out with a negative effect of possessing relatively less of \( Z_{id} \). These challenges operate in reverse where the bias results in overestimation.

The second term in equation (8) also shows when the bias will be greatest. The magnitude of the bias increases in \( \gamma \) because the compensating differential plays an important role in affecting district outcomes, and similarly increases in \( \alpha \) because more of the compensating differential is required to ensure close elections. The bias also increases in \( \frac{\sigma^2}{\sigma^2 \varepsilon} \), as noise becomes relatively less important than candidate characteristics in determining whether an election is close. The influence of \( \beta \) on the size of bias is ambiguous because a greater impact of compensating differentials on candidate vote shares reduces the importance of noise in generating close elections but also reduces the size of the compensating differential needed to overcome the difference in \( X_{id} \).

### 3.4 The bias of PCRD designs in general

To demonstrate the general applicability of the bias, I now relax the functional form and distributional assumptions imposed on \( V_{id} \) and \( Y_{id} \). First, candidate \( i \)'s vote share is now an unrestricted random function \( v(X_{id}, X_{jd}, Z_{id}, Z_{jd}, \epsilon_{id}, \epsilon_{jd}) \), where the randomness comes from the realization of two independent shocks \( (\epsilon_{id}, \epsilon_{jd}) \) that are independent of all other variables such that \( v(X_{id}, X_{jd}, Z_{id}, Z_{jd}, \epsilon', \epsilon'') \) is fixed for any given \( (\epsilon', \epsilon'') \). Second, I assume that potential outcomes are additively separable between \( X_{id} \) and \( Z_{id} \), but otherwise allow flexible effects of these variables to vary across districts:

**Assumption 3.** \( y(X_{id}, Z_{id}) = \tau_{d} X_{id} + g(Z_{id}) \).

Additive separability excludes the possibility that the effect of \( X_{id} \) varies with compensating differ-
entials $Z_{id}$. This assumption is not necessary for bias to emerge in PCRD designs, but facilitates a simple bias decomposition.

The following proposition establishes the bias of the PCRD estimator:

**Proposition 2.** Under Assumptions 1(a) and 3, the bias of the PCRD estimator is:

$$\mathbb{E}[\hat{\tau}_{PCRD}] - \tau_{PCRD} = \mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1].$$

This result demonstrates that PCRD designs generally yield biased estimates of the LATE of $X_{id}$. By not imposing Assumption 2, bias can emerge from characteristics that are unconditionally correlated with $X_{id}$ and/or correlations with $Z_{id}$ that are induced or altered by the need for compensating differentials for elections to be tied.

Extending the intuitions from the example with a single compensating differential, the following proposition establishes three sufficient conditions for PCRD designs to yield unbiased estimates of the LATE of $X_{id}$:

**Proposition 3.** Under Assumptions 1(a) and 3, the PCRD estimator is unbiased if one of the following conditions holds at the point of discontinuity:

(i) Assumption 2 holds, $V_{id} \perp X_{id}, X_{jd}$, and the function $v$ is injective in $Z_{id}$ for any $(Z_{jd}, \epsilon_{id}, \epsilon_{jd})$;

(ii) $g(z) = g(z')$ for all $z \neq z'$; or

(iii) Whenever $V_{id} = V_{jd}$, $v(1,0,z,\epsilon_{id},\epsilon_{jd}) = v(0,1,z,\epsilon_{jd},\epsilon_{id})$ for $\epsilon_{id} \neq \epsilon_{jd}$ and any $Z_{id} = Z_{jd} = z$.

Condition (i) amounts to ensuring that candidate vote shares are not affected by, or correlated with, $X_{id}$ and $X_{jd}$; this means that compensating differentials are not required for elections to be close. Condition (ii) allows for compensating differentials to counteract the electoral advantage of $X_{id}$, but requires that they not affect district outcomes. This condition includes two cases: no compensating differential affects $Y_{d}$, or the net effect of different compensating differentials exactly cancels out.
Condition (iii) similarly ensures that noise—factors that benefit a candidate that are not features of the candidate themselves—exclusively compensates for the electoral advantage of $X_{id}$.

4 Implications for applied research

The preceding analysis showed that PCRD designs seeking to isolate effects of an elected politician characteristic risk inducing post-treatment biases. This section explores the implications for applied research, starting by considering the viability of imposing one of the conditions in Proposition 3. Since these conditions are unlikely to hold in many applications, I then consider strategies to mitigate the threat to internal validity. I finally discuss the implications of redefining treatment to encompass $X_{id}$ and all compensating differentials.

4.1 Imposing an additional assumption

Perhaps the most appealing method for addressing the concern that PCRD estimates may be biased is to explicitly invoke and empirically substantiate one of the conditions in Proposition 3. Researchers could assume, at the discontinuity, that $X_{id}$ does not affect candidate vote shares (condition (i)), that compensating differentials induced by variation in $X_{id}$ do not affect the outcome of interest (condition (ii)), or that compensating differentials are not required because random electoral shocks counterbalance compensating differentials (condition (iii)). I focus on conditions (i) and (ii), given that condition (iii) is similar to condition (ii) in claiming that other factors affecting election results do not shape post-election district outcomes and because this knife-edge condition is unlikely to hold when condition (i) does not also hold. Unfortunately, neither assumption is easily validated and both assumptions conflict with theories of voting behavior.

4.1.1 Empirical challenges

Compellingly showing that no compensating differential affects the outcome of interest is particularly difficult. First, where compensating differentials are observed, strong support for this
condition requires that a researcher further show that each compensating differential has no effect on the outcome around the discontinuity. Finding identification strategies for each potential compensating differential—or even just the most plausible compensating differentials—is unlikely to be feasible. Moreover, if the effects of compensating differentials are heterogeneous, this challenge is exacerbated by the need for these estimates to be local to close elections. Second, because compensating differentials like candidate competence can be difficult to accurately measure, researchers can rarely be confident that unobserved compensating differentials are not affecting the outcome.

Validating that characteristic $X_{id}$ does not affect candidate vote shares around the discontinuity is more attainable. Since $X_{id}$ and $V_{id}$ are both observed, a single test showing that $X_{id}$ does not affect $V_{id}$ among candidates around the discontinuity would suffice. A compelling test showing that candidate $i$’s gender or party affiliation does not affect their vote share requires an additional research design—such as reforms encouraging certain candidate types or evidence from a conjoint experiment—exogenously varying $X_{id}$. A less compelling test might instead show a limited correlation between $X_{id}$ and $V_{id}$. However, researchers cannot simply use the ex post sample of close elections where $X_{id}$ and $V_{id}$ are uncorrelated by construction. To capture the local effect of $X_{id}$ among candidates that end up in close races, researchers might estimate conditional treatment effects in a subset of the data where elections are predicted to be close.

4.1.2 Theoretical challenges

Compounding the empirical challenge of validating conditions (i) and (ii), both assumptions are often theoretically implausible. Researchers using PCRD designs are usually interested in characteristics like gender or partisan affiliation because they expect these characteristics to impact outcomes that voters likely also care about. For $X_{id}$ not to influence candidate vote shares—as condition (i) requires—when it does affect district outcomes, voters would need to be oblivious to, or not vote on the basis of, the characteristic’s expected impact on the outcome of interest to the researcher or other outcomes voters are concerned about. For example, Gagliarducci and Paserman
(2012) argue that Italian voters are unlikely to vote on the basis of government termination risks and are poorly informed about whether gender might precipitate termination. However, termination risk could correlate with other outcomes that determine vote choice and voters attribute to gender. For condition (ii), the existence of compensating differentials that do not affect the district outcome of interest would require voters to wrongly believe that these characteristics affect an outcome they care about or not care about the outcome of interest to the researcher.

Although there is debate about the extent to which voters vote “rationally,” most theories of voting behavior suggest that at least some voters observe candidate characteristics and are aware of how such characteristics affect outcomes they care about. Even where electorates are only partially informed about the link between characteristics and outcomes, candidates with identical vote shares should produce identical welfare outcomes—broadly construed to encompass any outcome that matters to voters which different candidates could affect—in expectation if limited information is aggregated efficiently (e.g. Fowler 2018). Equal vote shares could reflect equally effective candidates committing to policies that converge on the median voter’s preferred policy (Downs 1957) or comparative advantages of one candidate on some dimensions being counteracted by the comparative advantages of other candidates on other dimensions (Krasa and Polborn 2010). The latter explanation does not imply that characteristic $X_{id}$ has no effect on a given outcome of interest, just that other characteristics produce offsetting effects on other outcomes that leave the median voter indifferent between the two candidates. Even if voters are not fully rational, the possibility that easily-observed candidate characteristics that impact important outcomes are simultaneously uncorrelated with candidate vote shares is implausible in many contexts.

4.2 Mitigating threats to internal validity

Where neither condition (i) or condition (ii) can be sustained, PCRD designs lack the design-based foundation for isolating the effect of characteristic $X_{id}$ because candidate that narrowly win are expected to differ in other consequential ways too. Although the post-treatment bias of PCRD designs is likely smaller than for designs that do not restriction attention to comparisons between
relatively similar candidates (see section SI.4), I next discuss three potential strategies to mitigate the threat to the internal validity of PCRD estimates and their limitations.

4.2.1 Continuity tests

As Table 1 illustrates, most studies using PCRDs conduct continuity—or balance—tests to validate that potentially-confounding district- or candidate-level characteristics do not vary discontinuously at the point of discontinuity. This typically entails estimating $\lim_{v \downarrow v} E[Z_{idk}|\Delta_d = v] - \lim_{v \uparrow v} E[Z_{0dk}|\Delta_d = v]$ to test the null hypothesis that $E[Z_{idk}|\Delta_d = 0, X_{id} = 1, X_{jd} = 0] - E[Z_{idk}|\Delta_d = 0, X_{id} = 0, X_{jd} = 1] = 0$ for each observable covariate $k$; for example, Anastasopoulos (2016), Brollo and Troiano (2016), and Gagliarducci and Paserman (2012) show that men and women in close races are similar in terms of education, experience, campaign donations, or ideology.

In standard RD designs, finding only differences consistent with statistical chance corroborates Assumption 1 (Imbens and Lemieux 2008). However, if neither condition (i) nor condition (ii) is compelling, researchers using PCRD designs should now expect to observe differences in candidate-level compensating differentials. In contrast, district-level characteristics remain useful for balance tests though, because the assumptions underpinning PCRD designs imply continuity in district-level characteristics. This is because district-level characteristics can determine the types of races that are close and the degree to which compensating differentials are required by affecting electoral advantages, but cannot vary across candidate types within a given race.\footnote{For example, $V_{id}$ clearly does not depend on $Z_{id} = Z_{jd} = Z_d$ in equation (5). However, a district-level characteristic could \textit{differentially} affect candidate type $X_{id} = 1$, e.g. if $V_{id} = \alpha X_{id} Z_d + \beta_1 (Z_{id} - Z_{jd}) + \beta_2 (X_{id} - X_{jd}) Z_d + \epsilon_{id} - \epsilon_{jd}$. Nevertheless, the effect of $Z_d$ on $Y_d$—given by $E[Y_d|\alpha + \beta_1 (Z_{d1} - Z_{jd1}) + \beta_2 Z_d + \epsilon_{id} - \epsilon_{jd} = 0]$—is the same when types $X_{id} = 1$ and $X_{id} = 0$ narrowly win, and thus cancels out. The magnitude of $Z_{d1} - Z_{jd1}$ required to compensate for $\alpha + \beta_2 Z_{d2}$, rather than $\alpha$, increases though.}

Because PCRD designs imply discontinuities in some candidate-level covariates where conditions (i) and (ii) are implausible, continuity tests operate differently. Whereas failing to reject continuity in observable covariates is desirable in standard RD designs, detecting discontinuities in theoretically-plausible compensating differentials in PCRD designs can instead serve as a manipulation check that guides researcher interpretation of PCRD estimates. As I discuss below, character-
izing compensating differentials—and thus potential forms of bias—can inform efforts to bounding estimates or reinterpret treatment. Conversely, observing few meaningful differences is consistent both with conditions (i) or (iii) holding—an implicit claim in studies documenting no significant discontinuity in observable candidate-level characteristics in support of their PCRD design—and with unobserved compensating differentials, a lack of statistical power, or even spurious PCRD estimates.

4.2.2 Covariate adjustment

Where the assumptions necessary for identification do not obviously hold, a strategy common to many statistical approaches is covariate adjustment. This involves adjusting, or “controlling,” for predetermined potential confounders to the greatest extent possible using observable covariates. Gagliarducci and Paserman (2012) address the imbalances between men and women that narrowly won mayoral elections in Italy by adjusting for various covariates, including those on which significant imbalances are observed. Covariate adjustment could be implemented by regressing $Y_d$ on $T_d$, a subset of (demeaned) compensating differentials $Z_{id}^{\text{cond}} \subset Z_{id}$, and $T_d Z_{id}^{\text{cond}}$.

However, covariate adjustment does little to address the post-treatment bias that arises in PCRD designs. Since equation (7) must always hold, covariate adjustment does not increase the plausibility of condition (i) because conditioning on $Z_{id}^{\text{cond}}$ induces or accentuates the need for compensating differentials in terms of other covariates $Z_{id} \setminus Z_{id}^{\text{cond}}$ that are not adjusted for. For example, a researcher using a PCRD design to estimate effects of electing university-educated politicians might condition on ideology because they are concerned that better-educated politicians are in close races because they espouse unpopular policy positions. Even if covariate adjustment breaks the correlation between education and ideology, university-educated politicians in close races with non-university-educated politicians with similar ideologies must still, in expectation, differ in other ways—such as non-educational dimensions of competence—to remain in close races. Covariate adjustment can increase the plausibility of conditions (ii) and (iii) by increasing the share of variation in candidate vote share explained by noise, i.e. reducing signal to noise ratio $\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2}$. However,
removing bias still requires that researchers accurately measure and fully adjust for all relevant compensating differentials, including compensating differentials induced by conditioning, that affect district-levels outcomes.

4.2.3 Bounding and correcting effect magnitudes

Where none of the additional conditions that guarantee unbiased PCRD estimates are plausible, a more promising strategy—at least for more limited researcher objectives—is to use $\hat{\tau}_{PCRD}$ and the preceding theoretical structure to bound the effect of $X_{id}$ or correct estimates of its effects. While such strategies cannot remove bias, they can help establish the direction of the effect or the type and degree of bias due to compensating differentials that would nullify a given estimate.

The preceding discussion of when PCRD designs over and underestimate the effect of $X_{id}$ illustrates the benefits and drawbacks to bounding. Indeed, underestimation—which allows the researcher to claim that the effect of $X_{id}$ is not smaller in magnitude than $\hat{\tau}_{PCRD}$—occurs when the combined effect of compensating differentials affect the outcome in the same way as $X_{id}$ at the discontinuity. This relatively strong conclusion for non-null findings could be substantiated by using continuity tests to identify compensating differentials and then providing theoretical or empirical evidence to argue that the effects of $X_{id}$ and $Z_{id}$ operate in the same direction at the discontinuity. For example, if primary voters are averse to ideological extremists and such candidates must compensate by being more competent on average, then Hall’s (2015) findings might understate the general election penalty associated with selecting extreme candidates in primary elections. However, establishing the direction of an effect is harder when $\hat{\tau}_{PCRD}$ fails to reject the null hypothesis because we cannot be sure if underestimation accounts for accepting the null.

In the spirit of Rosenbaum (2002), bias correction is also possible where plausible estimates of the effect of compensating differentials on district outcomes can be imputed. If $g(Z_{id})$ is linear and $\hat{\gamma}_k$ is a reliable estimate of the LATE of each compensating differential $Z_{idk}$ at the point of discontinuity, then Proposition 2 implies that the PCRD estimate could be corrected to obtain the
Latent of $X_{id}$ as follows:\footnote{Letting $Z_{id}$ include higher-order polynomials and interactions between characteristics, the Weierstrass approximation theorem ensures that $\sum k \gamma_k Z_{id}^k$ can approximate any $g(Z_{id})$.}

\[ \hat{\tau}_{PCRD}^{corr} = \hat{\tau}_{PCRD} - \sum_k \hat{\gamma}_k \hat{\mu}_k. \]  

where each $\hat{\mu}_k$ estimates $E[Z_{idk}|\Delta_d = 0, X_{id} = 1, X_{jd} = 0] - E[Z_{idk}|\Delta_d = 0, X_{id} = 0, X_{jd} = 1]$ using a (dis)continuity test for observable compensating differentials. Even without reliable estimates of $\gamma_k$, researchers could examine the sensitivity of their results to plausible values of $\gamma_k$. In the gender example, this could involve estimating the difference in candidate competence between men and women that narrowly win and then examining the implications for the estimates when $\gamma_k$ represents a 0.25, 0.5, 0.75, and 1 standard deviation effects. Section S1.5 provides examples of how attention to theorized mechanisms can inform $\gamma_k$.

### 4.3 Expanding the conception of treatment

Given the challenges of isolating effects of politician characteristic $X_{id}$, an alternative approach is to redefine the target estimand by expanding the conception of treatment to include compensating differentials. In particular, researchers might explicitly focus on a compound treatment incorporating effects of both $X_{id}$ and all compensating differentials induced or altered by PCRD designs (and their interactions) at the point of discontinuity. Hall (2015:24) adopts this type of approach when noting that his PCRD estimate of the effect of selecting ideological extremists in U.S. primary elections on general election outcomes “includes the component of the overall effect that comes from the change in ideology, but also includes any other factors that differ between the two types of candidates.” He thus distinguishes between a specific individual characteristic—extremism—and the bundle of correlated characteristics characterizing a typical extremist. This logic extends to distinguishing the effects of electing a candidate representing the positions of the Democratic party from electing candidates that are Democrats. In the context of women winning primary elections, Bucchianeri (2018:445) similarly defines his estimand as the “causal effect of nominating a female...
candidate, not the causal effect of gender,” and explicitly notes that this bundle could include compensating differentials that ensure women remain in close races with men. The key advantage of redefining the treatment of interest to include correlated characteristics is that PCRDs designs can now yield unbiased estimates, albeit for a different estimand.

Formally, this reconceptualization entails focusing on joint potential outcomes $Y_d(X_d, Z_d)$, where $Z_d$ are the characteristics of the winning candidate in district $d$. The following proposition characterizes the compound treatment effect identified by the PCRD estimator under the standard RD continuity assumption:

**Proposition 4.** Under Assumption 1(a), the PCRD estimator identifies:

$$
E[\hat{\tau}_{PCRD}] = \int y(1, z) f_{Z|\Delta=0, X=1, X_{jd}=0} d\zeta - \int y(0, z) f_{Z|\Delta=0, X=0, X_{jd}=1} d\zeta,
$$

where $f_{Z|\cdot}(\cdot)$ denotes the conditional probability density function of $Z_{id}$.

Where the distribution of $Z_{id}$ differs across candidates of type $X_{id} = 1$ and $X_{id} = 0$ in close elections, this result shows that $\hat{\tau}_{PCRD}$ captures effects of both $X_{id}$ and $Z_{id}$. In Hall’s (2015) case, $X_{id}$ represents ideological extremism and $Z_{id}$ captures all other characteristics—both those that naturally bundle with $X_{id}$ and those induced, accentuated, or attenuated by conditioning on close elections that are affected by $X_{id}$—of extremists, such as being less likely to be women. Moreover, because many permutations of $(Z_{id}, Z_{jd})$ can produce close elections, winning candidates of type $X_{id}$ can experience many values of $Z_{id}$. The compound treatment effect estimated by PCRD designs then captures the weighted average of the effects due to differences in $Z_{id}$ across elected politicians of different types that won close elections. Where Assumption 1(a) holds, bias cannot arise because all variables that differ at the point of discontinuity are now considered part of the treatment.

Reconceptualizing potential confounders as part of the effect identified by PCRD designs suggests a different role for continuity tests. While continuity tests for district-level covariates still serve to validate Assumption 1(a), candidate-level covariate tests yield estimates of $\hat{\mu}_k$ that can now help to characterize the average comparison captured by $\hat{\tau}_{PCRD}$ in Proposition 4. Substantial
differences in $Z_{1dk} - Z_{0dk}$ at the point of discontinuity suggest that $Z_{idk}$ may be an important component of the compound treatment, whereas the reverse holds for covariates where discontinuities are not found. Hall (2015) adopts such an approach by examining whether extremist primary winners differ from non-extremist winners on other dimensions in his analysis of mechanisms.

There are, however, three drawbacks to broadening the notion of treatment; the importance of these trade-offs will vary by application. First, it is hard to fully characterize treatment in many empirical applications because compensating differentials should generally exist but continuity tests can lack the statistical power to detect differences in $Z_{id}$ (Hartman 2020) and researchers may struggle to adequately measure relevant elements of $Z_{id}$. Beyond the label “$X$ and all its compensating differentials,” PCRDs may then lack clarity about what defines the bundle of characteristics that distinguish different types of elected candidate.

Second, as Proposition 4 shows, the external validity and interpretability of PCRD estimates may be limited because the design is unlikely to capture typical or constant bundles of characteristics. Rather than capture characteristics that are unconditionally correlated with possessing $X_{id}$, PCRD designs identify compound treatments defined by the correlations between characteristics that exist after conditioning on close elections where $X_{id}$ affect vote shares. Consequently, extremists that win close races against non-extremists may be atypical of extremists in general and even atypical of extremists that narrowly win any type of primary election. Moreover, treatments are likely to be heterogeneous across candidates—a violation of SUTVA—since different combinations of compensating differentials can generate close elections. For example, some extremists that narrowly win may be relatively more competent and others may offer more appealing platforms to overcome any penalty associated with their ideological extremism.

Third, bundled treatments limit the degree to which PCRD designs can test specific theories or inform certain policy decisions. Theories often specify “all else equal” comparative statics for different variables that PCRD designs cannot distinguish because all candidate-level characteristics, albeit to differing degrees, are considered part of the compound treatment. For example, PCRD designs cannot reveal whether ideological extremists lose general elections because of their policy
positions, differences in competence between extremist and non-extremist candidates that narrowly win, or a combination of both that cannot be pinned down.

The extent to which this inability to distinguish the contribution of different elements of the compound treatment limits the relevance of PCRD estimates to policymakers likely depends on the policy tools available. On one hand, policymakers with constrained choice sets may not need to know which part of the treatment matters, only that policies that encourage (or discourage) politicians of type $X_{id}$ to run for office or help such candidates win elections should be favored. For example, local party committees might alter candidacy rules to avoid selecting extremist candidates that lose general elections. On the other, the limited information conveyed by PCRD estimates is less helpful where policymakers are picking between policies that encourage candidates of type $X_{id}$ rather than type $Z_{idk}$. Understanding the mechanism driving the PCRD estimates could be consequential for reformers seeking to understand whether they should adopt gender quotas or require more specific competencies of their candidates.

5 Conclusions

This article shows that PCRD designs—an increasingly popular approach used to estimate effects of a specific characteristic, or bundle of characteristics, of elected politicians on downstream outcomes—generally require imposing substantially stronger assumptions than standard RD designs. This is because the treatment variable in this non-standard RD application is defined both by winning close elections and a candidate characteristic that can affect selection into the sample of narrow candidate winners of different types. I have shown that such post-treatment conditioning causes PCRD estimates to capture the effect of the specific characteristic of interest together with all the compensating differentials required for candidates with the characteristic of interest to still be in close races.

PCRD designs thus yield biased estimates of the LATE exclusively attributable to the characteristic that defines treatment, except under two additional assumptions: (i) the elected candidate
characteristic of interest does not affect the candidate’s vote share; or (ii) no compensating differential affects the outcome of interest. Unfortunately, neither sufficient condition for unbiased estimates is plausible in many contexts and both difficult are to test. Accordingly, PCRD designs often lack the capacity to pinpoint how electing different types of politician impact outcomes including political representation, accountability, and participation.

Although the credibility and feasibility of different responses is study-specific, researchers can attempt to combat this challenge in several different ways. One approach is to mitigate threats to external validity by using theory and data to bound treatment effects. Another approach broadens the definition of treatment to exclude the possibility of candidate-level confounding by defining the estimand of interest to include both the characteristic of interest and all compensating differentials. Both approaches entail trade-offs, either in terms of internal validity or conceptualization of treatment, but could nevertheless yield compelling—if more limited—inferrnces in certain settings.
References


Supplementary Information

Contents

SI.1 Published articles using PCRD designs .................................................. SI2
SI.2 Proof of propositions ................................................................................ SI2
SI.3 Example with two compensating differentials ........................................ SI5
SI.4 Do PCRD designs reduce bias relative to observational designs? .......... SI7
SI.5 Differential empirical implications of theorized mechanisms ............... SI8
Published articles using PCRD designs

Table S1 lists 78 published articles using PCRD designs.

Proof of propositions

**Proof of Proposition 1:** The following equalities prove the result:

\[
E[\tilde{\tau}_{RD}] = \lim_{v \downarrow 0} E[Y_{id}|\Delta_{id} = v] - \lim_{v \uparrow 0} E[Y_{id}|\Delta_{id} = v] \\
= \lim_{v \downarrow 0} E[Y_{id}(1)|\Delta_{id} = v] - \lim_{v \uparrow 0} E[Y_{id}(0)|\Delta_{id} = v] \\
= E[Y_{id}(1)|\Delta_{id} = 0] - E[Y_{id}(0)|\Delta_{id} = 0] \\
= E[Y_{id}(1) - Y_{id}(0)|\Delta_{id} = 0] \\
= \tau_{RD},
\]

where the first equality follows from Lebesgue’s Convergence Theorem (since \(E[Y_{id}|\Delta_{id} = v]\) is finite), the second equality follows from the definition of potential outcomes, the third equality follows from Assumption 1, and the fourth equality uses linearity of expectations to reformulate the expression as the causal effect of \(T_{id}\).

**Proof of Proposition 2:** The following derivation proves the result:

\[
E[\tilde{\tau}_{PCRD}] = \lim_{v \downarrow 0} E[Y_{d}|\Delta_{d} = v] - \lim_{v \uparrow 0} E[Y_{d}|\Delta_{d} = v] \\
= \lim_{v \downarrow 0} E[Y_{id}(1)|\Delta_{id} = v, X_{id} = 1, X_{jd} = 0] - \lim_{v \uparrow 0} E[Y_{id}(1)|\Delta_{id} = v, X_{id} = 0, X_{jd} = 1] \\
= E[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - E[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] \\
= E[\tau_{d} + g(Z_{id}) + \nu_{d}|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - E[g(Z_{id}) + \nu_{d}|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] \\
= E[\tau_{d}|\Delta_{d} = 0] + E[g(Z_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - E[g(Z_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] \\
= \tau_{PCRD} + E[g(Z_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - E[g(Z_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1],
\]
Table SI1: Studies using PCRD designs to estimate effects of winning candidate characteristics

<table>
<thead>
<tr>
<th>Candidate characteristic</th>
<th>Published articles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ideology</strong></td>
<td>Hall (2015), Hall and Thompson (2018).</td>
</tr>
<tr>
<td><strong>Race, ethnicity, religion, or clan</strong></td>
<td>Beach and Jones (2017), Bhalotra et al. (2014), Hopkins and McCabe (2012), Vogl (2014), Xu and Yao (2015).</td>
</tr>
</tbody>
</table>

*Note:* This list, which excludes unpublished studies and may be incomplete, is a lower bound on this design’s usage.
where the first equality follows from Lebesgue’s Convergence Theorem (since \( \mathbb{E}[Y_d|\Delta_d = v] \) is finite), the second equality follows from the definition of potential outcomes and the definition of \( \Delta_d \), the third equality follows from Assumption 1(a), the fourth equality substitutes the structural definition of potential outcomes in Assumption 3, the fifth equality uses the linearity of the expectation operator and independence of \( v_d \) from all other variables, and the sixth equality follows from

\[
\tau_{\text{PCRD}} = \mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0] = \mathbb{E}[\tau_d|\Delta_d = 0].
\]

**Proof of Proposition 3:** To demonstrate sufficiency, it suffices to show that each condition implies that \( b_{\text{PCRD}} := \mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] = 0 \) in the result given by Proposition 2.

For condition (i): given Assumption 2 and, by the assumption in condition (i), that the function \( v \) is independent of \( X_{id} \) and \( X_{jd} \), \( v \) can be written as \( v(Z_{id}, Z_{jd}, \varepsilon_{id}, \varepsilon_{jd}) \). Since the assumption in condition (i) further requires that \( v \) is injective in \( Z_{id} \) for any given \( (Z_{jd}, \varepsilon_{id}, \varepsilon_{jd}) \), \( v(Z_{1d}, Z_{jd}, \varepsilon_{id}, \varepsilon_{jd}) = v(Z_{0d}, Z_{jd}, \varepsilon_{id}, \varepsilon_{jd}) \) implies \( Z_{1d} = Z_{0d} \). Consequently, \( g(Z_{1d}) = g(Z_{0d}) \), and thus \( b_{\text{PCRD}} = 0 \).

For condition (ii): the condition implies that \( \mathbb{E}[g(Z_{id})|\Delta_{id} = 0] = \mathbb{E}[g(Z'_{id})|\Delta_{id} = 0] \) for all \( Z_{id}, Z'_{id} \), and thus \( b_{\text{PCRD}} = 0 \) holds.

For condition (iii): the condition ensures that \( Z_{id} = Z_{jd} \) whenever \( \Delta_{id} = 0 \). This implies that

\[
\mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1] = \mathbb{E}[g(Z_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0]
\]

whenever \( \Delta_{id} = 0 \), and thus \( b_{\text{PCRD}} = 0 \).

**Proof of Proposition 4:** The expectation of \( \hat{\tau}_{\text{PCRD}} \) is given by:

\[
\mathbb{E}[\hat{\tau}_{\text{PCRD}}] = \lim_{\nu \uparrow 0} \mathbb{E}[Y_d|\Delta_d = v] - \lim_{\nu \uparrow 0} \mathbb{E}[Y_d|\Delta_d = v]
\]

\[
= \lim_{\nu \uparrow 0} \mathbb{E}[Y_{id}(1)|\Delta_{id} = v, X_{id} = 1, X_{jd} = 0] - \lim_{\nu \uparrow 0} \mathbb{E}[Y_{id}(1)|\Delta_{id} = v, X_{id} = 0, X_{jd} = 1]
\]

\[
= \mathbb{E}[y(1, Z_{id}) + \nu_d|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[y(0, Z_{id}) + \nu_d|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1]
\]

\[
= \int y(1, z)f_{Z_{id}|\Delta_{id}=0, X_{id}=1, X_{jd}=0}(z)dz - \int y(0, z)f_{Z_{id}|\Delta_{id}=0, X_{id}=0, X_{jd}=1}(z)dz
\]
where the first equality follows from Lebesgue’s Convergence Theorem (since \( E[Y_d|\Delta_d = v] \) is finite), the second equality follows from the definition of potential outcomes and the definition of \( \Delta_d \), the third equality follows from Assumption 1(a), the fourth equality substitutes the structural definition of potential outcomes, and the fifth equality applies the definition of conditional independence and the independence of \( \upsilon_d \) from all other variables.

**SI.3 Example with two compensating differentials**

While the main paper considered the simplest example with a single compensating differential, similar insights obtain for multiple compensating differentials. Consider an example with three relevant candidate characteristics: the characteristic of interest, \( X_{id} \in \{0,1\} \); and two potential compensating differentials, \( Z_{id} = (Z_{id1}, Z_{id2}) \). Two compensating differentials allows for the possibility that compensating differentials may be correlated and act in opposing ways.

I assume that \( Z_{id1} - Z_{jd1} \) and \( Z_{id2} - Z_{jd2} \) are both normally distributed across candidates according to \( N(0,1) \), with a correlation between these differences of \( \rho \in (-1,1) \). By normalizing the standard deviations of the differences in candidate characteristics to 1, \( \beta_k \) and \( \delta_k \) can be interpreted as standardized effects. The difference between the idiosyncratic vote share shocks \( \epsilon_{id} - \epsilon_{jd} \) is independently and normally distributed across districts according to \( N(0,\sigma^2_{\epsilon}) \). I impose the following functional forms on the vote and outcome functions:

\[
V_{id} = \alpha \frac{X_{id} - X_{jd}}{2} + \sum_{k=1,2} \beta_k \frac{Z_{idk} - Z_{jdk}}{2} + \frac{\epsilon_{id} - \epsilon_{jd}}{2},
\]

\[
y(X_{id}, Z_{id}) = \tau X_{id} + \sum_{k=1,2} \gamma_k Z_{idk}
\]

The following result characterizes the bias of the PCRD estimator in this case:

**Proposition SI1.** Where \( Z_{idk} - Z_{jdk} \sim N(0,1) \) for each \( k = 1,2 \), \( \text{Corr}[Z_{id1} - Z_{jd1}, Z_{id2} - Z_{jd2}] = \rho \), and \( \epsilon_{id} - \epsilon_{jd} \sim N(0,\sigma^2_{\epsilon}) \), the bias of PCRD estimator is:

\[
\mathbb{E}[\hat{\tau}_{PCRD}] - \tau_{PCRD} = -\sum_{k=1,2} \gamma_k \left( \frac{\alpha(\beta_k + \rho \beta_{-k})}{\beta_1^2 + \beta_2^2 + 2\rho \beta_1 \beta_2 + \sigma^2_{\epsilon}} \right).
\]

SI5
If $\rho$ is not too negative, the compensating differentials cause the PCRD estimator to underestimate (overestimate) the magnitude of $X_{id}$'s effect when the sign of $\tau_{PCRD}$ agrees (disagrees) with the signs of $\gamma_1$ and $\gamma_2$. If $\rho$ is sufficiently negative, the PCRD estimator overestimates the magnitude of $X_{id}$'s effect when $|\gamma_1 - \gamma_2|$ is sufficiently large.

Proof of Proposition SI1. Given that $Z_{id1} - Z_{jd1}, Z_{id2} - Z_{jd2}$, and $\varepsilon_{jd} - \varepsilon_{id}$ are normally distributed as stated in the proposition:

$$
E[Z_{dk} - Z_{jdk}|\Delta_d] = E[Z_{idk} - Z_{jdk}] + \frac{\text{Cov}[Z_{idk} - Z_{jdk}, \Delta_d]}{\text{Var}[\Delta_d]}(\Delta_d - E[\Delta_d]),
$$

where $E[Z_{idk} - Z_{jdk}] = 0$ and $E[\Delta_d] = \alpha$ (because $Z_{id1} - Z_{jd1}, Z_{id2} - Z_{jd2}$, and $\varepsilon_{jd} - \varepsilon_{id}$ are all centered on 0), and $\text{Var}[\Delta_d] = \beta_1^2 + \beta_2^2 + 2\rho\beta_1\beta_2 + \sigma^2_{\varepsilon}$ (due to the independence of $\varepsilon_{id} - \varepsilon_{jd}$ from $Z_{idk}$ and the distributional assumptions on $Z_{id}$), recalling that $\Delta_d = \alpha + \sum_{k=1}^2 \beta_k(Z_{1dk} - Z_{0dk}) + \varepsilon_{1d} - \varepsilon_{0d}$. We then obtain the following expressions by conditioning on $\Delta_d = 0, X_{idk} = 1$, and $X_{jdk} = 0$:

$$
E[Z_{id1} - Z_{0d1}|\Delta_d = 0, X_{id} = 1, X_{jd} = 0] = -\frac{\alpha(\beta_1 + \beta_2\rho)}{\beta_1^2 + \beta_2^2 + 2\rho\beta_1\beta_2 + \sigma^2_{\varepsilon}},
$$

$$
E[Z_{id2} - Z_{0d2}|\Delta_d = 0, X_{id} = 1, X_{jd} = 0] = -\frac{\alpha(\beta_2 + \beta_1\rho)}{\beta_1^2 + \beta_2^2 + 2\rho\beta_1\beta_2 + \sigma^2_{\varepsilon}},
$$

where the covariance terms reflect independence of $\varepsilon_{id} - \varepsilon_{jd}$ from $Z_{idk} - Z_{jdk}$ and the distributional assumptions on $Z_{idk} - Z_{jdk}$. Applying Proposition 2 then yields the result in Proposition SI1.

It is obvious from inspection of this result that, when $\rho$ is not too negative and $E[\tau_d|\Delta_d = 0] > 0$, $X_{id}$'s effect is underestimated (overestimated) when $\gamma > (\prec)0$. The reverse holds when $E[\tau_d|\Delta_d = 0] < 0$. When $\rho$ is sufficiently negative and $\gamma_k$ share the same sign ($\gamma_k \geq 0$, without loss of generality), it is possible that $\beta_1\gamma_1 + \beta_2\gamma_2 + \rho(\beta_1\gamma_2 + \beta_2\gamma_1) < 0$, and thus overestimation occurs when $E[\tau_d|\Delta_d = 0] > 0$. At the lower bound of $\rho = -1$, the condition holds when $(\beta_1 - \beta_2)(\gamma_1 - \gamma_2) < 0$. ■

The additional insight that emerges from this example with multiple compensating differentials
is that PCRD estimates can also be unbiased in the knife-edge case where conflicting effects of
the compensating differentials on the outcome exactly cancel out. Such a scenario would only
hold if voters systematically failed to appreciate the implications of at least one compensating
differential for district-level outcomes. For example, competent candidates might be less popular
despite competence improving post-election outcomes in their district.

SI.4 Do PCRD designs reduce bias relative to observational designs?

Since no bias mitigation strategy is perfect, it is natural to wonder whether the biases introduced by
PCRD designs are greater than alternative research designs. To illuminate this question, I compare
the PCRD design to a difference in means design within the general framework of section 3.4. The
difference in means estimator is given by $\hat{\tau}_{DM} := \mathbb{E}[Y_{id}|\Delta_d > 0, X_{id} = 1] - \mathbb{E}[Y_{id}|\Delta_d > 0, X_{id} = 0]$, where this comparison of winners of different types need not restrict attention to races where the
top two candidates differed in their type $X_{id}$.\(^1\) This design, which admits observations from races
that are not close, estimates effects of $X_{id}$ away from $\Delta_d = 0$.

The following proposition establishes when $\hat{\tau}_{PCRD}$ is less biased—relative to its target estimand—
than $\hat{\tau}_{DM}$:

**Proposition SI2.** Under Assumptions 1(a) and 3, $\hat{\tau}_{PCRD}$ is less biased than $\hat{\tau}_{DM}$ when:

$$
\left| \mathbb{E}[g(Z_{id})|\Delta_d = 0, X_{id} = 1, X_{jd} = 0] - \mathbb{E}[g(Z_{id})|\Delta_d = 0, X_{id} = 0, X_{jd} = 1] \right| \\
\leq \left| \mathbb{E}[g(Z_{id})|\Delta_d > 0] - \mathbb{E}[g(Z_{id})|\Delta_d \leq 0] \right|.
$$

**Proof of Proposition SI2:** The magnitude of the bias on the left hand side follows directly from

\(^1\)More complex conditioning strategies, such as selection on observables or difference-in-differences, could be
used instead.
Proposition 2. The bias of $\hat{\tau}_{DM}$ is derived as follows:

$$E[\hat{\tau}_{DM}] = E[Y_{id}|\Delta_{id} > 0, X_{id} = 1] - E[Y_{id}|\Delta_{id} > 0, X_{id} = 0]$$

$$= E[Y_{id}(1)|\Delta_{id} > 0, X_{id} = 1] - E[Y_{id}(1)|\Delta_{id} > 0, X_{id} = 0]$$

$$= E[\tau_d|\Delta_{id} > 0] + E[g(Z_{id}) + \nu_d|\Delta_{id} > 0, X_{id} = 1] - E[g(Z_{id}) + \nu_d|\Delta_{id} > 0, X_{id} = 0],$$

$$= E[\tau_d|\Delta_{id} > 0] + E[g(Z_{id})|\Delta_{id} > 0, X_{id} = 1] - E[g(Z_{id})|\Delta_{id} > 0, X_{id} = 0],$$

where the first line follows from consistent estimation of the conditional means, the second follows from the individual-level definition of potential outcomes, the third line follows from Assumption 3, and the fourth line follows from the independence of $\nu_d$ from all other variables. The bias of the difference in means estimator is $E[\hat{\tau}_{DM}] - E[\tau_d|\Delta_{id} > 0]$. ■

The inequality indicates that PCRD design are less biased when differences in outcomes due to differences in $Z_{id}$ at the point of discontinuity are smaller than the corresponding differences between candidates of different types that won with any vote share. While it is plausible that candidates with the same vote share are more similar, it is ultimately an empirical question whether more popular candidates that differ in terms of $X_{id}$ are more similar in terms of $Z_{id}$.

SI.5 Differential empirical implications of theorized mechanisms

Beyond or in addition to bounding exercises, researchers can attempt to separate compound treatments by testing implications that help to distinguish which candidate characteristics drive PCRD estimates. The goal of this approach is thus to substantiate a form of the claim that compensating differentials do not affect the outcome of interest. Observing meaningful discontinuities among other variables may again be useful in guiding the components of the compound treatment that researchers should focus attention on. It should be emphasized that analyses of mechanisms rely on strong assumptions (Bullock, Green and Ha 2010), and are most feasible when there are few compensating differentials to distinguish the effects of characteristic $X_{id}$ from.

Researchers could provide evidence consistent with some components of the compound treat-
ment driving an effect, and not others, in at least two ways. A first approach uses an underlying theory to identify post-treatment variables that are expected to change if $X_{id}$ drives the effect at the discontinuity, but are not expected to change if relevant compensating differentials instead drive the effect. For example, Gagliarducci and Paserman (2012) argue that municipal governments led by women are more likely to be terminated early due to resistance from lower-ranked men on the municipal council. If the PCRD estimate of the effect of gender on early termination were confounded by women that win close elections possessing different levels of competence than men that win close elections, then we should expect the treatment to influence government performance outcomes as well. The lack of such evidence in Gagliarducci and Paserman (2012) suggests that competence is not driving the PCRD estimates, either because it is not a compensating differential or because competence does not affect local government outcomes in the study’s specific context.

A second approach based on moderation leverages subgroup variation where there are theoretical or contextual reasons to believe that some components of a compound treatment are more likely to be activated in certain subgroups than other components. For example, Gagliarducci and Paserman (2012) argue that the effect of electing a woman on early government termination should be greater in parts of Italy where unfavorable attitudes toward women are greater and there was limited prior history of women in office, whereas the downstream effect of electing a more competent candidate may not vary with such baseline conditions. However, it should be cautioned that definitively attributing differences across subgroups to a particular covariate also requires exogenous variation in that covariate.