CEDE
Centro de Estudios sobre Desarrollo Económico

Addressing Bias in Politician Characteristic Regression Discontinuity Designs

Documento CEDE

Estudiantes PEG

#24

Santiago Torres

Agosto de 2023

Universidad de los Andes
Colombia
Facultad de Economía

Serie Documentos Cede, 2023-24 ISSN 1657-7191 Edición electrónica. Agosto de 2023

infocede@uniandes.edu.co
http://economia.uniandes.edu.co

Impreso en Colombia – Printed in Colombia

CEDE
Centro de Estudios sobre Desarrollo Económico

Documento CEDE

Estudiantes PEG

Descripción: estas publicaciones contienen los resultados de las tesis de maestría (PEG o MEcA) de los estudiantes de la Facultad que se consideran merecedoras de este honor. Extraordinariamente, por postulación de al menos un profesor de planta de la Facultad, se pueden considerar trabajos de estudiantes de los programas de la Facultad. En este caso, la decisión sobre su publicación o no dependerá de la dirección del CEDE con el aporte del concepto de al menos un profesor de planta (distinto al que postula el trabajo). Como, contrario a las tesis, estos documentos no necesariamente han acreditado el cumplimiento con los requisitos de ética en la investigación, al presentar el documento se debe diligenciar el formato para ese efecto.

Universidad de los Andes
Colombia
Facultad de Economía

Addressing Bias in Politician Characteristic Regression Discontinuity Designs*

Santiago Torres†

August 23, 2023

Abstract

Politician characteristic regression discontinuity (PCRD) designs are a popular strategy when attempting to casually link a specific trait of an elected politician with a given outcome. However, recent research has revealed that this methodology often fails to retrieve the target causal effect—a problem also known as the PCRD estimation bias. In this paper, I provide a new econometric framework to address this limitation in applied research. First, I propose a covariate-adjusted local polynomial estimator that corrects for the PCRD estimation bias provided all relevant confounders are observed. I then leverage the statistical properties of this estimator to propose several decompositions of the bias term and discuss their potential applications. Next, I devise a strategy to assess the robustness of the new estimator to omitted confounders that could potentially invalidate results. Finally, I illustrate these methods through an application: a PCRD aimed at evaluating the impact of female leadership during the COVID-19 pandemic.

Keywords: Regression discontinuity designs, Close elections, Bias correction, Sensitivity analysis.
JEL Classification: C18, C51, P00

*Foremost, I want to thank Leopoldo Fergusson for his dedicated mentorship, well-intentioned advice, and friendship during the five years working with him at Los Andes. Additionally, I want to express my gratitude to Manuel Fernández for his guidance and support as co-advisor of this project. Adriana Camacho and John Marshall, who served on my dissertation jury, provided insightful feedback and suggestions. Finally, I want to thank Valentina Castilla and Paula Torres for their comments on the early versions of this project.

†Pearson Institute Pre-Doctoral Scholar, Harris School of Public Policy, University of Chicago. 1307 East 60th Street, Chicago, IL 60637.
Master's degree candidate, Universidad de los Andes, Facultad de Economía, Cra 1 No 18A - 12, Bogotá, Colombia, s.torresp1@uniandes.edu.co

Tratamiento del sesgo de estimación en los diseños de regresión discontinua con características de políticos.*

Santiago Torres**

23 de agosto de 2023

Resumen

Los diseños de regresión discontinua basados en características de los políticos (PCRD, por sus siglas en inglés) son una estrategia popular cuando se intenta vincular casualmente un rasgo específico de un político electo con un resultado determinado. Sin embargo, investigaciones recientes han mostrado que esta metodología a menudo no recupera el efecto causal objetivo, un problema también conocido como sesgo de estimación en los diseños de PCRD. En este artículo, proporciono un nuevo marco econométrico para abordar esta limitación en la investigación aplicada. En primer lugar, propongo un estimador de polinomios locales ajustado por covariables que corrige el sesgo de estimación PCRD siempre que se observen todos los factores de confusión relevantes. A continuación, aprovecho las propiedades estadísticas de este estimador para proponer varias descomposiciones del término de sesgo y discutir sus posibles aplicaciones. A continuación, diseño una estrategia para evaluar la robustez del nuevo estimador frente a variables omitidas que podrían invalidar los resultados. Por último, ilustro estos métodos mediante una aplicación: un PCRD destinado a evaluar el impacto del liderazgo femenino durante la pandemia COVID-19.

Palabras clave: Diseños de regresión discontinua, Elecciones Reñidas, Corrección del Sesgo, Análisis de sensibilidad.
Clasificación JEL: C18, C51, P00

*Ante todo, quiero agradecer a Leopoldo Fergusson su dedicada mentoría, sus bienintencionados consejos y su amistad durante los cinco años que trabajé con él en Los Andes. Además, quiero expresar mi gratitud a Manuel Fernández por su orientación y apoyo como coasesor de este proyecto. Adriana Camacho y John Marshall, que formaron parte del jurado de mi tesis, me proporcionaron comentarios y sugerencias muy enriquecedores. Por último, quiero dar las gracias a Valentina Castilla y Paula Torres por sus comentarios en las primeras versiones de este proyecto.

**Pasante predoctoral en el Pearson Institute, Harris School of Public Policy, University of Chicago. 1307 East 60th Street, Chicago, IL 60637.
Candidato a maestro en economía, Universidad de los Andes, Facultad de Economía, Cra 1 No 18A - 12, Bogotá, Colombia, s.torresp1@uniandes.edu.co

1 Introduction

The question of how particular electoral outcomes influence other social, economic, and political events is central to many social sciences. However, isolating causal links between these topics is challenging, given how intricately these topics are related. A common strategy to uncover these causal relationships relies on estimating regression discontinuity (RD) designs comparing the outcomes of several tight electoral contests (Lee, 2008). Close elections may provide a credible situation in which candidates or parties that garner just enough support to win may be similar, and thus comparable, to candidates or parties that narrowly lose. Thereby, RD designs can identify (local) causal effects under some assumption guaranteeing the comparability between winners and losers (Cattaneo, Idrobo, & Titiunik, 2019).

A popular variant of close election RD designs is known as the Politician Characteristic Regression Discontinuity (PCRD) design[1]. PCRD designs seek to determine the impact of a politician's characteristic X (such as their gender, race, party affiliation, etc.) on a given outcome Y. In essence, this approach compares instances where a politician who possesses the interest trait narrowly defeats a rival who does not, to elections where the opposite occurs. Therefore, PCRD designs do not contrast narrow winners to narrow losers, as formulated in the canonical close election RD design. Instead, they compare outcomes in places where a narrow winner features a trait to those same outcomes in places where the narrow winner lacks the trait (equivalently, where the narrow loser has the trait).

However, Marshall (2022) demonstrates that PCRD designs typically fail to recover the desired causal effect due to the existence of confounding factors. Even though PCRD designs can help hold constituency-level confounders constant and causally identify outcome differences between places with and without an appointed politician with characteristic X, these disparities may not exclusively reflect differences in the highlighted trait. One reason is simple and amply recognized: politicians who exhibit trait X may consistently display other personal traits that can also account for changes in the outcome. However, the lesser-known aggravating problem is that such differences between winners and losers may *unintendedly* and *unavoidably* arise when conditioning the sample to close races.

For example, using a PCRD design to study the impact of women on politics in places with gender discrimination may overlook that women who can compete with men in close races are likely to be exceptionally talented or that their male opponents are relatively incompetent. Hence, if talent or competence matters for outcomes, the PCRD design will unavoidably estimate a compound treatment comprising the effect of gender and all distinctive traits of female contestants involved in close elections. In a broader sense, this distortion can be thought of as a type of post-treatment bias (Montgomery, Nyhan, & Torres, 2018),

[1] Following Marshall (2022), at least 126 articles using a PCRD design have been published in recent years, often in prestigious journals and recurrently cited: "These articles have consistently appeared in prominent journals in political science and economics: 21% were published in the American Journal of Political Science, American Political Science Review, or Journal of Politics, whereas 5% were published in the American Economic Review, Econometrica, the Quarterly Journal of Economics, or the Review of Economic Studies. According to Google Scholar, these studies had collectively amassed 11,774 citations by March 8, 2022." (p.5)

insofar as it emanates from conditioning on a variable (vote-margin) that is impacted by the treatment (the trait of interest). The resulting estimation bias is thus a major concern for researchers and policymakers who rely on this identification strategy to quantify the impact of trait X on a particular outcome.

This paper develops the first widely applicable econometric framework for bounding, decomposing, and eliminating PCRD estimation bias in applied research. First, I prove that, provided all relevant confounders are observed, a local polynomial estimator with a linear-in-parameters covariate adjustment can consistently remove the PCRD estimation bias. This method is fully non-parametric, meaning its effectiveness does not rely on assuming a specific relationship between the outcome and a politician's characteristics. Moreover, the resulting estimator can be a more precise alternative to conventional techniques, given it can simultaneously reduce estimation bias and variance in multiple empirical settings. In addition, the proposed estimator is asymptotically normal, and its convergence rate is that of univariate local polynomial regressions regardless of the number or nature of the confounders.

Second, I leverage the estimator's linear structure to construct a simple and independent estimator of the PCRD estimation bias. I then show that this estimator can be decomposed into several economically meaningful quantities. The primary goal of this breakdown is to determine the extent to which compensating differentials or other ancillary characteristics explain the results in these locations. To that end, I propose three distinct decompositions: one that assesses the influence of each observable trait on the compound effect; a Kitagawa-Oaxaca-Blinder decomposition that evaluates whether the bias is primarily due to differences in individual attribute endowments or their differential impact on the outcome; and a detailed version that combines both of these breakdowns.

While the proposed estimator is general enough to correct bias in any PCRD design, its effectiveness depends on the ability of a researcher to control for *all* relevant confounders. Whenever this is impossible, which unfortunately is almost always the case, the covariate-adjusted estimator is not guaranteed to completely eliminate the PCRD estimation bias. This circumstance is thus comparable to the omitted variables bias (OVB) in conventional linear regressions. As a result, I leverage my estimator's linearity to adapt existing OVB sensibility methods to assess the credibility of PCRD designs in a more realistic situation in which some individual characteristics may not be incorporated in the correction. These methods provide reasonable bounds on the PCRD bias based on the observed data so that a researcher can determine how likely it is that the omission of a relevant confounder could potentially invalidate results.

As a running example, I use the work of Bruce, Cavgias, Meloni, and Remígio (2022), who evaluate the impact of female mayors on epidemiological outcomes in Brazil during the COVID-19 pandemic using a PCRD design. This application is particularly relevant since there is empirical evidence of several compensating differentials emerging from close elections. Therefore, absent an assessment of the role of compensating differentials, this calls into question the causal interpretation of the traditional PCRD design conducted by the authors. My method offers such tools, and in this specific application, I show that observed differences in individual attributes do not significantly alter the authors' results. Moreover, I can credibly discard the existence of any unobserved confounder that could drastically change the study's results. Hence, I show that my methodology provides researchers with tools to assess the credibility of causal effect estimates resulting from PCRD designs even when the estimation bias cannot be dismissed.

The present study is most closely aligned with the work of Marshall (2022), who first characterizes the PCRD estimation bias. While Marshall's primary contribution was highlighting the existence and origins of the PCRD bias, he also explored potential solutions. However, his suggestions require strong parametric assumptions about the bias structure or the use of meta-analysis to infer the sign or size of the bias. While my proposal pursues the same type of solutions—covariate adjustment and bounding— it introduces a rigorous and systematic method for capping and correcting estimation bias in all PCRD designs. In that sense, it furnishes a more general and less discretionary method than Marshall's approach.

Due to its applicability, this paper is relevant to all empirical works that use PCRD designs as an identification strategy. These include the literature emphasizing the role of ascriptive characteristics such as gender (Bhalotra & Clots-Figueras, 2014; Clots-Figueras, 2011; Ferreira & Gyourko, 2014), race (Hopkins & McCabe, 2012), ethnicity (Beach & Jones, 2017) , clan (Xu & Yao, 2015), or religious identity (Bhalotra, Clots-Figueras, Cassan, & Iyer, 2014) in other downstream outcomes. My results also pertain to research focusing on the role of politicians' past like its prior incumbency (Lewis, Nguyen, & Hendrawan, 2020), criminal history (Chemin, 2012), and seniority (Fowler & Hall, 2015). Likewise, my findings apply to studies stressing the importance of party membership (Pettersson-Lidbom, 2008), ideology (Hall, 2015), and partisan alignment with other levels of the government (Bracco, Lockwood, Porcelli, & Redoano, 2015).

This paper is also related to the emerging field of research on selection bias correction and covariate adjustment in RD designs. A key contribution in this area is the work of Frölich and Huber (2019a), who developed a non-parametric kernel method for estimating local average treatment effects in RD designs with *selection on observables*. Similar to their work, I demonstrate that it is possible to construct local polynomial estimators with linear-in-parameters covariate adjustments that can asymptotically remove selection bias. However, the relative advantage of my approach is that my estimator can eliminate selection bias while preserving a simple linear structure that is useful for other analyses. Likewise, my estimator is also related to the work of Calonico, Cattaneo, Farrell, and Titiunik (2019a), who explored linear-in-parameters covariate adjustments in RD designs to reduce variance. The main difference is that my paper investigates how these specifications can restore identification when additional confounders are present. Other studies have proposed eliminating bias originated by potential confounders by re-weighting Nadaraya-Watson estimators using inverse propensity scores (Peng & Ning, 2019). My proposal deviates from this approach since it uses local regression methods for the correction. Lastly, Mukherjee, Banerjee, and Ritov (2021) extend standard RD designs by modeling the running variable as a linear function of observed covariates, yielding semiparametric estimators capable of correcting bias. In contrast, I consider a scenario in which selection may be correlated with both the treatment and the forcing variable.

My work also speaks to research concerning omitted variable bias sensitivity analysis. I contribute to this literature by being the first paper to propose OVB sensitivity analysis tailored to local polynomial estimators and RD designs. Furthermore, it is also the first approach to implement sensitivity analysis in the presence of both unobserved confounding and post-treatment bias. My proposal is based on Cinelli and Hazlett (2020a), who created a set of tools for sensitivity analysis in regression models that are valid regardless of the functional form of the treatment assignment mechanism or the distribution of unobserved confounders. I build on this approach since it is flexible enough to adapt to RD design estimation and overcomes several interpretation challenges arising in conventional approaches in economics like those of Altonji, Elder, and Taber (2005) and Oster (2019)[2]. Other works on this subject that relate to my work include Imbens (2003), Hosman, Hansen, and Holland (2010) and Rosenbaum and Rubin (1983).

The remainder of this paper is organized as follows. Section 2 motivates the origins and consequences of the PCRD estimation bias following the theoretical framework described in Marshall (2022), and introduces relevant mathematical notation. Section 3 proposes a method for eliminating the PCRD bias if all relevant confounders are observed using a covariate-adjusted polynomial estimator. Moreover, it describes several decompositions of the bias estimator that can be useful for applied research and public policy. Section 4 introduces the sensitivity analysis to account for possible omitted variables. Section 5 illustrates the practical implications of my methodological framework using as an example the work of Bruce et al. (2022). Section 6 concludes.

2 The politician characteristic regression discontinuity design estimation bias

In this section, I will briefly discuss the causes of the PCRD estimation bias and offer a mathematical formulation to assist in its comprehension. I largely adopt the notation and terminology used in Marshall (2022) to create a link between the two papers.

The PCRD estimation bias originates from the inability to distinguish between a politician's trait X and other attributes that may also impact an outcome variable Y. Following Marshall (2022), the PCRD bias will likely emerge from two sources. On the one hand, it is highly unlikely that the characteristic X is randomly assigned. That means candidates with feature X may systematically differ from politicians without X in other personal traits. Therefore, since these disparities are systematic, they will also manifest when comparing candidates running in close elections. For example, candidates from politically underrepresented groups are more likely to have less experience in public office than their peers. As a result, disparities in outcomes claimed to be caused by a politician's ethnicity, gender, or religion might as well be explained by their lack of experience in public office. Since many of these ancillary characteristics may be unobserved or difficult to measure, disentangling the effect of X from other ascriptive characteristics constitutes a *con-*

[2] For a complete discussion on the matter, refer to Section 6.3 of Cinelli and Hazlett (2020a).

founding problem. As in observational studies, being unable to isolate the contribution of the trait of interest on the outcome via the RD design ultimately results in an estimation bias.

On the other hand, by focusing only on close elections, PCRD designs are likely to compare non-representative samples of politicians, given only candidates with singular characteristics may end up in tight electoral contests. This happens because *conditioning* in a narrow victory event is far from uninformative: if the votes that a candidate obtains are largely determined by the trait X, then being in a close electoral competition *necessarily requires* that other attributes must be unbalanced for both candidates to receive roughly the same number of votes. These induced differences are referred to as *compensating differentials*, and their influence on estimation can be viewed as a form of *post-treatment bias*[3]. Conceptually, this is a different source of bias from the confounding problem since it doesn't fade even if trait X is as-good-as-randomly assigned. In realistic scenarios, practitioners should expect PCRD estimation bias to arise from both confounding and compensating differentials, potentially interacting in a way that estimation bias exceeds that of an observational study contrasting the naive difference in outcomes between candidates of different types.

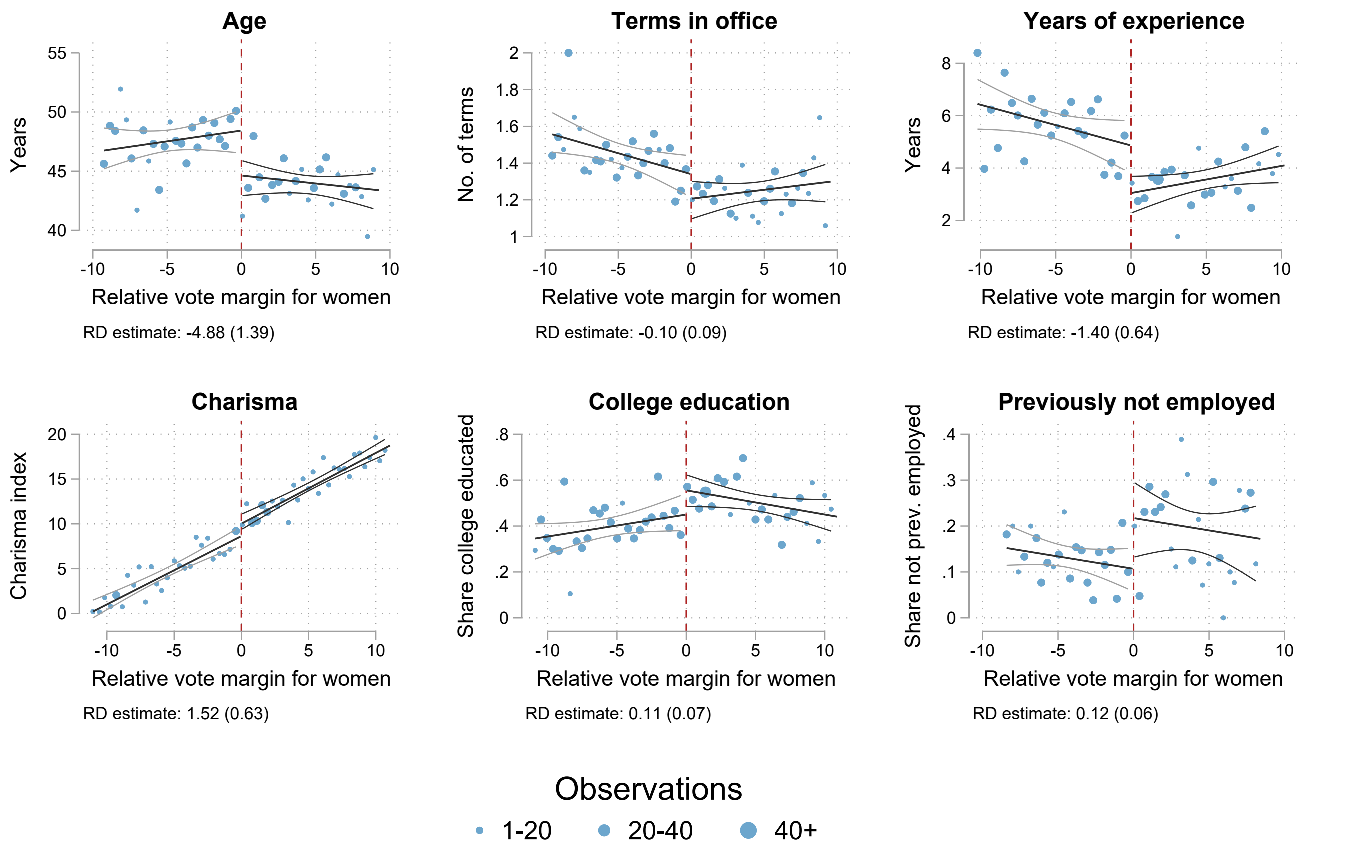
The intuition behind compensating differentials may be best understood within the context of an example. There is a large body of literature suggesting the existence of a gender bias in politics[4], which postulates that the average voter prefers male leaders over female leaders because they believe male political leaders are more capable in public office. As a result, if gender bias considerably influences voting decisions, a contest between two candidates who are identical except for their gender will result in the male winning by a large margin. Therefore, the fact that a male and a female finish in a close race suggests that something else is *compensating* for the adversely perceived trait of being a woman. For example, women who end up in close races may be considerably more educated and prepared than men. Alternatively, men in tight electoral contests may be less popular than the average male politician.

For instance, Gagliarducci and Paserman (2012) leverage PCRD design to investigate whether the gender of the chief executive affects the likelihood of the elected administration surviving until the end of its mandate in a parliamentary system. Figure 1 empirically confirms that women involved in the close races used for this study are systematically different from male candidates. For one, many attributes are systematically correlated with gender. In particular, female mayors elected by a narrow margin are disproportionately less experienced, younger, and more likely to have never held a job. These facts are consistent with the historical underrepresentation of women in the workforce and politics. For another, there is also suggestive evidence that female candidates probably need to compensate with other positive traits to compete closely with men. To see this, notice how elected women politicians in these races have significantly more education

[3]Post-treatment bias, also known as endogenous selection (Elwert & Winship, 2014) or collider bias (Pearl, 2009), is the distortion in the estimation of an average treatment effect resulting from including variables in estimation that are causally affected by the treatment. Researchers may induce post-treatment bias in two ways: by conditioning the sample based on a post-treatment variable or by controlling for one. Post-treatment bias can pose a threat to identification even in completely experimental designs. For further information, please refer to the works of Acharya, Blackwell, and Sen (2016); Montgomery et al. (2018); Robins (2023), and references therein.

[4]Some evidence of this phenomenon can be found in Lawless (2015).

FIGURE 1: *Compensating differentials in Gagliarducci and Paserman (2012)*



Notes: RD plots of several individual characteristics of elected politicians in Italian municipal governments between 1993 and 2003 that competed in close races involving candidates of different genders. The $X-$axis accounts for the difference in votes between the most voted women and the most voted men in the contest. Bins were selected following the mimicking variance evenly-spaced method developed by Calonico, Cattaneo, and Titiunik (2015) using spacings estimators. Local linear polynomial estimates are reported following Calonico, Cattaneo, and Titiunik (2014) using a triangular kernel and an MSE-optimal bandwidth, together with standard errors clustered at the municipality level. Data retrieved from Gagliarducci and Paserman (2012).

and are more charismatic than their male counterparts.

Consequently, PCRD designs identify not only the differences in outcome caused by trait X, but also differences caused by other characteristics associated with having trait X in a close election. Unfortunately, the latter represents a bias term that may misrepresent the size and sign of the target treatment effect. For instance, Gagliarducci and Paserman (2012) argue that women have a higher probability of early resignation than men based on their gender. But Figure 1 shows this finding could also be (partially) explained by women being comparatively less experienced and younger than their male counterparts. The above example demonstrates how conventional PCRD designs frequently fail to achieve the desired causal effect unless additional strong assumptions are made (Marshall, 2022, Proposition 3). Given this limitation, researchers face two options. One possibility is that they acknowledge the difficulty of retrieving the desired effect and interpret PCRD estimates as capturing both effects of X and all compensating differentials. Alternatively, they can attempt to produce bounds or eliminate the estimation bias using additional information. This paper centers its contribution on the possibility of the second idea.

2.1 The formal setup

Electoral district-level data is usually employed for estimating close election RD designs. Let X_{id} represent a binary random variable, with $X_{id} = 1$ indicating that candidate i in district d has trait X. A pair of politicians with the value of X will be referred to as being of the same type. Additionally, each candidate also has a set of m accompanying characteristics $\mathbf{Z}_{id} = (Z_{id1}, \cdots, Z_{idm})$ that are both correlated with vote levels and an interest district-level variable Y_d. We will restrict to districts where the top two candidates are not the same type.

Let V_{1d} and V_{0d} respectively denote the vote shares of the most popular politician of type $X_{id} = 1$ and $X_{id} = 0$ in district d. For the sake of simplicity, candidate 1 will always denote the politician who possesses the interest attribute (as in $X_{id} = 1$), whereas candidate 0 represents the politician who does not have the feature. By considering only each district's winner and runner-up candidate, the variable $\Delta_d = V_{1d} - V_{0d}$ reflects the percentage point difference in vote shares between the politician with $X_{id} = 1$ and the one without. Thereby, this variable can act as a running variable for a sharp RD design to the extent that it perfectly discriminates whenever a politician of type 1 wins the election over a candidate of type 0:

$$X_d := \begin{cases} 1 & \text{if } \Delta_d \geq 0, \\ 0 & \text{if } \Delta_d < 0. \end{cases}$$

Since observing X_d can be interpreted as a treatment status, the most natural way to think about causal effects is in terms of potential outcomes. More precisely, prospects can be thought of in two separate dimensions. The first is at the *individual level*, whenever candidate i wins and the outcome $Y_{id}(1)$ is observed, and the second is when it loses and the outcome $Y_{id}(0)$ occurs. The second is at the *district level*, provided either candidate 1 or 0 must be elected. These two observations justify the outcome variable being expressed as

$$Y_d = Y_d(X_d) = X_d Y_{1d}(1) + (1 - X_d) Y_{0d}(1)$$

A key feature of PCRD designs is that it *does not compare winners to losers, but winners of different types*. To appreciate the difference, consider a typical work aiming to determine incumbency advantages using an RD design. In such studies, the outcome variable is dummy variable Y_{id} informing whether individual i is elected in district d in the current elections. Hence, to determine whether winning previous elections affects the probability of reaching office, a researcher is interested in estimating the expected value of $Y_{id}(1) - Y_{id}(0)$. This is because this quantity reflects how much more probable *candidate i* is to win an election given that they held office in the prior electoral term compared to a scenario in which they did not. PCRD designs, in contrast, compare *two candidates* of opposing types. Moreover, the method cannot provide any information about $Y_{id}(0)$ insofar as comparisons are made *between winners*.

Likewise, the same logic applies to the appointed politician's accompanying characteristics $\mathbf{Z}_{id}$. Hence, these attributes admit a representation using potential outcomes as

$$\mathbf{Z}_d = \mathbf{Z}_d(X_d) = X_d \mathbf{Z}_{1d} + (1 - X_d) \mathbf{Z}_{0d}$$

Finally, note that the target causal effect can be described as

$$\tau_{PCRD} = \mathbb{E}[Y_d(1) - Y_d(0) | \Delta_d = 0],$$

which accounts for the expected difference in the outcome variable whenever a politician of type $X_{id} = 1$ wins the election by an infinitesimal margin over one of type $X_{id} = 0$.

2.2 Estimation using local polynomials

Because of their boundary properties, local polynomial estimators are the most frequently employed approach for estimating RD designs (Cheng, Fan, & Marron, 1997; Hahn, Todd, & Van der Klaauw, 2001). This method consists of two stages. First, the researcher must choose a bandwidth h around the cutoff where the estimation will occur. Next, practitioners must fit a local model $\tilde{Y}(\Delta_d, X_d; \theta)$, by estimating a parameter vector θ using as a weighting scheme a kernel function $K(\cdot)$[5]. The kernel function serves two purposes: it limits the estimation to observations whose running variable scores fall within the bandwidth while assigning more weight to units closer to a designated value. For PCRD designs, this threshold is 0, meaning that the local polynomial $\hat{\theta}$ estimator is derived from a weighted least squares problem given by

$$\hat{\theta}(h) = \underset{\theta}{argmin} \sum_d (Y_d - \tilde{Y}(\Delta_d, X_d; \theta))^2 K(\Delta_d / h)$$

As its name implies, local polynomial estimators result from requiring the local model $\tilde{Y}(\Delta_d, X_d; \theta)$ to be a polynomial of degree p. In practice, PCRD designs frequently feature (but are not restricted to) a local

[5] A kernel function is a non-negative real-valued integrable function that is 1) symmetric and 2) integrates to one over its support.

linear form such as[6]

$$\tilde{Y}_d = \tilde{Y}(\Delta_d, X_d; \tilde{\theta}) = \tilde{\alpha}_Y + X_d \tilde{\tau}_Y + \Delta_d \tilde{\beta}_{Y-} + (X_d \times \Delta_d)\tilde{\beta}_{Y+} \tag{2.1}$$

where $\tilde{\theta} = (\tilde{\alpha}_Y, \tilde{\tau}_Y, \tilde{\beta}_{Y-}, \tilde{\beta}_{Y+})' \in \mathbb{R}^4$ are the parameters to be estimated using weighted least squares. In particular, we are interested in $\tilde{\tau}_Y$ since it captures the effect that X_d has on $\tilde{Y}_d$ at the threshold $\Delta_d = 0$. It should be noted that all results in this paper are valid for any polynomial degree, despite the exposition concentrating on a linear specification. Moreover, a complete description of the estimation procedure for an arbitrary polynomial degree can be found in Appendix B.3.

Two sets of assumptions are required for these estimators to behave appropriately. On the one hand, some *regularity conditions* on the conditional expectation functions are needed to ensure consistency and asymptotic normality of the estimators. Appendix B.2 contains a detailed discussion of these requirements as offered by Calonico, Cattaneo, Farrell, and Titiunik (2019b). On the other hand, RD designs require an *identification assumption* for establishing causality, which is given by the following weak continuity assumption(Hahn et al., 2001).

Assumption 1 (Local continuity)**.** *The potential outcomes $Y_d(X_d)$ satisfy*

a) Continuity from above: $\lim_{\eta \downarrow 0} \mathbb{E}[Y_d(1)|\Delta_d = \eta] = \mathbb{E}[Y_d(1)|\Delta_d = 0]$

b) Continuity from below: $\lim_{\eta \uparrow 0} \mathbb{E}[Y_d(0)|\Delta_d = \eta] = \mathbb{E}[Y_d(0)|\Delta_d = 0]$

According to this assumption, the only way that potential outcomes can differ at the threshold is because a candidate of type X_d won the election. For this to hold true, the *districts* where a type 1 politician barely wins must be comparable to those they narrowly lose. This requirement is reasonable in close elections because random factors, such as the weather on election day, may determine several election outcomes whenever vote margins are tight[7].

Under Assumption 1, some regularity assumptions and a suitable choice of bandwidths sequences[8], the weighted least squares estimator $\hat{\tilde{\tau}}_Y$ of $\tilde{\tau}_Y$ identifies a causal effect

$$\begin{aligned}
\hat{\tilde{\tau}}_Y &= \lim_{\eta \downarrow 0} \mathbb{E}\,[\widehat{Y_d|\Delta_d = \eta}] - \lim_{\eta \uparrow 0} \mathbb{E}\,[\widehat{Y_d|\Delta_d = \eta}] \\
&\xrightarrow{\mathbb{P}} \lim_{\eta \downarrow 0} \mathbb{E}\,[Y_d|\Delta_d = \eta] - \lim_{\eta \uparrow 0} \mathbb{E}\,[Y_d|\Delta_d = \eta] \\
&= \mathbb{E}\,[Y_d(1)|\Delta_{id} = 0] - \mathbb{E}\,[Y_d(0)|\Delta_{id} = 0] \\
&= \mathbb{E}\left[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] - \mathbb{E}\left[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1\right],
\end{aligned}$$

[6]Following Gelman and Imbens (2019), researchers frequently limit to linear or quadratic specifications ($p = 1, 2$).

[7]Notably, the plausibility of this assumption has been thoroughly tested in practical scenarios with success (e.g. Eggers, Fowler, Hainmueller, Hall, and Snyder (2015); Hyytinen, Meriläinen, Saarimaa, Toivanen, and Tukiainen (2018)).

[8]Bandwidth sequences h_n should converge slowly to 0. More precisely, $h_n \to 0$ and $nh_n \to \infty$, where n denotes the sample size.

where X_{jd} is district d's runner-up candidate type and $\Delta_{id} = V_{id} - V_{jd}$ is the observed winner's margin. Notice that the second equality differs from the first in that it rewrites district-level outcomes Y_d in terms of candidate-level outcomes Y_{id}. Since PCRD always compares winners, we only observe $Y_{id}(1)$ and not $Y_{id}(0)$. Nevertheless, winners are of different types, thus changing the conditioning events. Since this is the standard method to estimate PCRD, I will frequently refer to $\hat{\hat{\tau}}_Y$ as $\hat{\tau}_{PCRD}$.

2.3 Motivating the PCRD estimation bias

The outcome variable Y_d in PCRD designs is likely to depend not only on the district's characteristics but also on the individual attributes of the elected politician. Therefore, the outcome variable should be a function of the studied trait X_{id}, as well as the other individual-level characteristics $\mathbf{Z}_{id}$. Conveniently, we can represent this dependence as

$$Y_{id}(1) = m(\Delta_{id}, \mathbf{Z}_{id}) + \varepsilon_d$$

where $m(\Delta_{id}, \mathbf{Z}_{id}) = \mathbb{E}[Y_{id}(1)|\Delta_{id}, \mathbf{Z}_{id}]$ is known as the associated regression function, and ε_d is an idiosyncratic error term satisfying $\mathbb{E}[\varepsilon_d|\Delta_{id}, \mathbf{Z}_{id}] = 0$. While in practice, $m(\cdot,\cdot)$ is an unknown arbitrary function, I will focus on a streamlined example to motivate how PCRD estimation bias emerges.

Assumption 2 (Simple data generation process)**.** *The data generating process of $Y_{id}(1)$ is*

$$Y_{id}(1) = \tau_d X_{id} + g(\mathbf{Z}_{id}) + \varepsilon_d \tag{2.2}$$

where $g : \mathbb{R}^m \to \mathbb{R}$ is some unknown smooth function.

Assumption 2 simplifies the expected theoretical relations between the outcome variable and the characteristics of the elected politician by assuming that $m(\Delta_{id}, \mathbf{Z}_{id}) = \tau_d X_{id} + g(\mathbf{Z}_{id})$. While this functional form facilitates exposition, it should be stressed that asymptotic bias will likely appear in PCRD designs with *arbitrary* data generation processes (Marshall, 2022, Proposition 4).

Under Assumption 2, traditional local linear estimators can fail to produce consistent estimators of τ_{PCRD}. To see this, note that

$$\begin{aligned}\hat{\tau}_{PCRD} &\xrightarrow{\mathbb{P}} \mathbb{E}\left[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] - \mathbb{E}\left[Y_{id}(1)|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1\right] \\ &= \tau_{PCRD} + \underbrace{\left(\mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] - \mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1\right]\right)}_{\text{PCRD estimation bias}}\end{aligned} \tag{2.3}$$

since $\tau_{PCRD} = \mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0] = \mathbb{E}[\tau_d|\Delta_d = 0]$. Therefore, the term

$$b_{PCRD} = \mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] - \mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id} = 0, X_{id} = 0, X_{jd} = 1\right]$$

will skew estimation whenever the influence of ancillary characteristics on the outcome varies with the candidate's type. This may occur when individual characteristics $\mathbf{Z}_{id}$ and type X_{id} are systematically correlated, or when conditioning on a close election leads to compensating differentials. Under any of these circumstances, $\hat{\tau}_{PCRD}$ identifies not only the direct effect that trait X has on the outcome, but also the influence of all other relevant characteristics disproportionately prevalent in either type in close elections.

Marshall (2022) made a significant contribution by demonstrating that the PCRD estimation bias represents a major threat to identification, given it is negligible only in a handful of razor-thin cases. To illustrate this, consider a scenario where only one additional factor Z_{id}, apart from the trait of interest, influences both the outcome Y_{id} and the number of votes V_{id} each candidate obtains. For simplicity, suppose that the votes cast in favor of a candidate depend on the relative endowment of ascriptive characteristics and on the imbalance of some idiosyncratic shocks (ξ_{id}, ξ_{jd}) with mean zero. For instance, we could model this dependence as

$$V_{id} = \zeta_1 \left(\frac{X_{id} - X_{jd}}{2} \right) + \zeta_2 \left(\frac{Z_{id} - Z_{jd}}{2} \right) + \left(\frac{\xi_{id} - \xi_{jd}}{2} \right)$$

where j denotes the characteristics of the opposing candidate. Under this assumption, the voting margin between the best-ranked candidate of type 1 and of type 0 in each district can be expressed as follows:

$$\Delta_d = V_{1d} - V_{0d} = \zeta_1 + \zeta_2 (Z_{1d} - Z_{0d}) + (\xi_{1d} - \xi_{0d})$$

Hence, the event of a close election ($\Delta_d = 0$) implies that

$$\underbrace{Z_{1d} - Z_{0d}}_{\text{Compensating differential}} = \frac{\xi_{0d} - \xi_{1d} - \zeta_1}{\zeta_2} \neq 0$$

this is, conditioning on a close race *requires* that $Z_{1d} \neq Z_{0d}$ in district d unless the accompanying term is exactly zero. Even when this may be the case for a particular district, on average, we would still observe that $\mathbb{E}[Z_{1d} | \Delta_d = 0] \neq \mathbb{E}[Z_{0d} | \Delta_d = 0]$, thus giving way to estimation bias.

From a broader perspective, conditioning on a close race is, by definition, sample selection based on a post-treatment variable, given that votes depend directly on trait X (the treatment). Hence, the bias caused by compensating differentials is a form of post-treatment bias. Specifically, because we ultimately end up comparing dissimilar groups, conditioning on a close race precludes the local comparability needed to retrieve causal effects via RD designs. More importantly, notice this pattern holds regardless of what Z_{id} and X_{id} are distributed and how they interrelate. For instance, PCRD estimation bias will arise even when X_{id} is independent of Z_{id}, as well as in a situation in which there are no systematic differences in Z_{id} between candidates of opposing types in the population (*i.e.,* $\mathbb{E}[Z_{1d} - Z_{0d}] = 0$).

3 PCRD estimation bias correction

Section 2 showed that, except in unlikely circumstances, conventional PCRD design estimations have asymptotic biases. Hence, according to Marshall (2022), a better tactic is to limit or correct the estimation bias rather than invoking unsuitable assumptions. In this section, I illustrate that when all of the politician characteristics that drive the bias are observed, a slight modification to the local polynomial estimator can accurately account for the bias under mild assumptions. Moreover, this adjusted estimator allows for several informative decompositions of the PCRD bias.

3.1 Motivation: Bias correction with a linear regression function

I begin by constructing a solution for the most basic scenario: the situation where Assumption 2 holds and $g(\cdot)$ is linear. This case was presented by Marshall (2022, Equation 9) as an example of how bounding and correcting effect magnitudes could strengthen the credibility of PCRD designs. Despite its apparent limitations, this case is useful as it effectively conveys most of the concepts needed to create a general solution. Consequently, I begin by considering a simplified data generation process.

Assumption 2′. *The data generating process of $Y_{id}(1)$ is*

$$Y_{id}(1) = \tau_d X_{id} + \sum_{l=1}^{m} \gamma_l Z_{idl} + \varepsilon_d; \quad \textit{for some } \gamma = (\gamma_1, \cdots, \gamma_l)' \in \mathbb{R}^m \tag{3.1}$$

By invoking Assumption 2′ and referring to equation (2.3), the bias takes a simple form:

$$\begin{aligned} b_{PCRD} &= \mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id}=0, X_{id}=1, X_{jd}=0\right] - \mathbb{E}\left[g(\mathbf{Z}_{id})|\Delta_{id}=0, X_{id}=0, X_{jd}=1\right] \\ &= \sum_{l=1}^{m} \gamma_l \left(\mathbb{E}\left[Z_{idl}|\Delta_{id}=0, X_{id}=1, X_{jd}=0\right] - \mathbb{E}\left[Z_{idl}|\Delta_{id}=0, X_{id}=0, X_{jd}=1\right]\right) \end{aligned}$$

Which means that

$$\hat{\tau}_{PCRD} = \hat{\hat{\tau}}_Y \xrightarrow{\mathbb{P}} \tau_{PCRD} + \sum_{l=1}^{m} \gamma_l \delta_l = \tau_{PCRD} + \delta'\gamma; \quad \delta = (\delta_1 \quad \delta_2 \quad \cdots \quad \delta_m)' \in \mathbb{R}^m$$

with

$$\delta_l = \delta_l^+ - \delta_l^- = \mathbb{E}\left[Z_{idl}|\Delta_{id}=0, X_{id}=1, X_{jd}=0\right] - \mathbb{E}\left[Z_{idl}|\Delta_{id}=0, X_{id}=0, X_{jd}=1\right]$$

In addition to being concise, the linear case clearly illustrates the nature of the estimation bias. In this setting, the total bias is the sum of distortions introduced by each ancillary attribute. Each of these deviations is the product of two elements. For one, δ_l accounts for the average difference in attribute Z_{idl} between politicians of type $X_{id}=1$ and type $X_{id}=0$ who competed in close races. In particular, δ_l might reflect compensating differentials induced by conditioning on tight electoral contests. Therefore, if gender is the interest trait, δ_l could represent how much more educated women are than men on average when they participate in close elections. For another, γ_l informs how relevant attribute Z_{idl} is to determining the outcome variable Y_d[9]. Hence, a simple mnemonic of the general bias structure is that every confounder distorts "its impact times its imbalance".

[9] More precisely, γ_l is the marginal effect on $m(\Delta_{id}, \mathbf{Z}_{id})$ of increasing a unit of Z_{idl}.

Furthermore, the linear structure makes bias correction relatively simple. In the light of Rosenbaum (2002) and Marshall (2022), if there existed consistent estimators $\hat{\hat{\delta}}_l$ and $\hat{\gamma}_l$ of δ_l and γ_l, then a consistent estimator $\hat{\tau}^{BC}_{PCRD}$ of τ_{PCRD} could be obtained through

$$\hat{\tau}^{BC}_{PCRD} = \hat{\tau}_{PCRD} - \sum_{l=1}^{m} \hat{\gamma}_l \hat{\hat{\delta}}_l \xrightarrow{\mathbb{P}} \tau_{PCRD}$$

Yet, is it possible to construct these estimators using only the information from the original design? On the one hand, consistent estimates of δ_l can potentially be retrieved under an additional assumption:

Assumption 3 (Local continuity of covariates)**.** *The potential outcomes $\mathbf{Z}_d(X_d)$ satisfy*

a) Continuity from above: $\lim_{\eta\downarrow 0}\mathbb{E}[\mathbf{Z}_d(1)|\Delta_d = \eta] = \mathbb{E}[\mathbf{Z}_d(1)|\Delta_d = 0]$

b) Continuity from below: $\lim_{\eta\uparrow 0}\mathbb{E}[\mathbf{Z}_d(0)|\Delta_d = \eta] = \mathbb{E}[\mathbf{Z}_d(0)|\Delta_d = 0]$

Like Assumption 1, Assumption 3 postulates that any observed discrepancies in the average traits of elected politicians in close races can only be attributed to the candidates' type. For this to be credible, close races should be able to rule out other district-level characteristics that can explain these discrepancies. For instance, if places where candidates of type 1 win by a small margin are systematically more urban than those in which politicians of type 0 triumph, then differences in education levels may be due to development levels rather than the politicians' type. Fortunately, in most empirical works, any credible scenario where Assumption 1 is true is also likely to constitute a situation where 3 is also valid.

Under Assumption 3, an RD design using local model (2.1) and as outcome variable X_dZ_{dl} leads to a local linear estimator satisfying

$$\hat{\hat{\tau}}_{X_dZ_{dl}} = \lim_{\eta\downarrow 0}\mathbb{E}[X_d\widehat{Z_{dl}|\Delta_d} = \eta] - \cancelto{0}{\lim_{\eta\uparrow 0}\mathbb{E}[X_d\widehat{Z_{dl}|\Delta_d} = \eta]} \xrightarrow{\mathbb{P}} \mathbb{E}\left[Z_{idl}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] = \delta_l^+$$

Likewise,

$$\hat{\hat{\tau}}_{(1-X_d)Z_{dl}} \xrightarrow{\mathbb{P}} \delta_l^-$$

so that

$$\hat{\hat{\tau}}_{X_dZ_{dl}} - \hat{\hat{\tau}}_{(1-X_d)Z_{dl}} \xrightarrow{\mathbb{P}} \delta_l^+ - \delta_l^- = \delta_l.$$

Hence, estimating traditional RD designs for every politician's trait Z_{idl} results in a consistent estimator of δ_l.

Nevertheless, even when consistent estimates of δ_l can be produced under reasonable assumptions, identifying parameter γ_l is more challenging. Specifically, the challenge lies in finding credible identification strategies that enable the separate identification of each parameter γ_l. Provided this limitation, Marshall (2022) suggests examining the sensitivity of their results to plausible values of γ_l. On the contrary, I show that, as long as all relevant covariates are observed, γ can be consistently estimated without the need for a separate identification strategy

Calonico et al. (2019a) discuss a covariate-adjusted local polynomial estimator to leverage additional information to produce a more efficient estimator in RD designs. Interestingly, for the approach to be effective, the covariates utilized for the adjustment must not vary at the threshold because doing so would introduce a bias to estimation. In the following paragraphs, I'll show how this estimator can successfully remove the bias from $\hat{\tau}_{PCRD}$. In essence, the idea is that using the covariate-adjusted estimator introduces a new bias into the PCRD estimator, which, if appropriately specified, may be able to cancel out the original PCRD bias.

In covariate-adjusted local polynomials estimators, rather than modeling the conditional expectation function locally using only the running variable as in (2.1), we want also to include the covariates in the functional form. There are several ways to accomplish this, but Calonico et al. (2019a) opt for a linear-in-parameters specification. In other words, this means that the canonical local linear model can be formulated as[10]

$$\dot{Y}_{id} = \dot{Y}(\Delta_d, X_d, \mathbf{Z}_{id}; \dot{\theta}) = \dot{\alpha}_Y + X_d \dot{\tau}_Y + \Delta_{id}\dot{\beta}_{Y-} + (X_d \times \Delta_d)\dot{\beta}_{Y+} + X_d \mathbf{Z}'_{id}\dot{\gamma}_+ + (1 - X_d)\mathbf{Z}'_{id}\dot{\gamma}_- \tag{3.2}$$

where $\dot{\theta} = (\dot{\alpha}_Y, \dot{\tau}_Y, \dot{\beta}_{Y-}, \dot{\beta}_{Y+}, \dot{\gamma}_+, \dot{\gamma}_-) \in \mathbb{R}^4 \times \mathbb{R}^m \times \mathbb{R}^m$ are the parameters to be estimated using weighted least squares. I will now concentrate on understanding the properties of the estimator $\hat{\dot{\tau}}_Y$ of $\dot{\tau}_Y$, as in model (2.1). To do so, I will first uncover a relationship between the new estimator $\hat{\dot{\tau}}_Y$ and the traditional estimator $\hat{\tau}_{PCRD} = \hat{\tilde{\tau}}_Y$.

Theorem 1. *Consider the local linear estimators* $(\hat{\dot{\alpha}}_Y, \hat{\dot{\tau}}_Y, \hat{\dot{\beta}}_{Y-}, \hat{\dot{\beta}}_{Y+}, \hat{\dot{\gamma}}_+, \hat{\dot{\gamma}}_-) \in \mathbb{R}^4 \times \mathbb{R}^m \times \mathbb{R}^m$ *resulting from an RD design using* (3.2) *as local model and* Y_d *as outcome variable. Then numerically*

$$\hat{\dot{\tau}}_Y = \hat{\tilde{\tau}}_Y - \left[(\hat{\tilde{\tau}}_{\mathbf{Z},+})'\hat{\dot{\gamma}}_+ - (\hat{\tilde{\tau}}_{\mathbf{Z},-})'\hat{\dot{\gamma}}_-\right] = \hat{\tau}_{PCRD} - \left[(\hat{\tilde{\tau}}_{\mathbf{Z},+})'\hat{\dot{\gamma}}_+ - (\hat{\tilde{\tau}}_{\mathbf{Z},-})'\hat{\dot{\gamma}}_-\right]$$

where

$$\hat{\tilde{\tau}}'_{\mathbf{Z},+} = \left(\hat{\tilde{\tau}}_{X_d Z_{d1}}, \cdots, \hat{\tilde{\tau}}_{X_d Z_{dm}}\right)_{1\times m}; \quad \hat{\tilde{\tau}}'_{\mathbf{Z},-} = \left(\hat{\tilde{\tau}}_{(1-X_d)Z_{d1}}, \cdots, \hat{\tilde{\tau}}_{(1-X_d)Z_{dm}}\right)_{1\times m}$$

Proof. This proposition is formulated assuming that a linear model (with respect to running variable) is used to specify local models (2.1) and (2.2). However, in Appendix B.4, I show how the equivalence sill holds for any local polynomial of degree p.

■

Theorem 1 illustrates the consequences of using the covariate-adjusted local polynomial estimator. Observe that applying a model like (3.2) results in an estimator that is identical to that of (2.1), except for the additional term $-[(\hat{\tilde{\tau}}_{\mathbf{Z},+})'\hat{\dot{\gamma}}_+ - (\hat{\tilde{\tau}}_{\mathbf{Z},-})'\hat{\dot{\gamma}}_-]$. As a result, whenever $\hat{\tau}_{PCRD}$ retrieves the desired effect in a particular RD design, including covariates introduces an undesirable bias term. Because of this, Calonico et al. (2019a) claim this particular covariate adjustment is useful only under conditions where this bias is zero.

[10]It is possible to include additional district-level covariates that do not vary in the threshold as in Calonico et al. (2019a) to reduce variance further.

However, I suggest that whenever the original estimator does not identify the target effect, this new term has the potential of correctly recentering the estimate. To assess this possibility, it is critical to determine what $\hat{\hat{\gamma}}_+$ and $\hat{\hat{\gamma}}_-$ capture.

Theorem 2. *Under Assumption 2′, regularity assumptions RA-1, RA-2, RA-3, and a suitable choice of bandwidths sequences,*

$$(\hat{\hat{\gamma}}_+, \hat{\hat{\gamma}}_-)' \xrightarrow{\mathbb{P}} (\gamma, \gamma)'$$

Proof. See Appendix B.5. ■

Despite being just nuisance parameters in the original formulation, Theorem 2 unveils that $\hat{\hat{\gamma}}_+$ and $\hat{\hat{\gamma}}_-$ are the estimators needed for bias correction whenever Assumption 2′ holds. Nonetheless, it should be stressed that this is possible only because *all* relevant confounders are observed. Without this assumption, notice that *i*) not all of the required "gammas" could be estimated, and *ii*) even those that could be estimated might be biased due to the omitted variable and post-treatment bias. Following this discussion, I summarize the main results of this section in the following theorem.

Theorem 3. *Under Assumptions 1 and 3, together with the hypothesis required for Theorem 2,*

$$\hat{b}_{PCRD}^{Linear} = (\hat{\hat{\tau}}_{\mathbf{Z},+})'\hat{\hat{\gamma}}_+ - (\hat{\hat{\tau}}_{\mathbf{Z},-})'\hat{\hat{\gamma}}_- \xrightarrow{\mathbb{P}} \delta'\gamma = b_{PCRD}$$

which means that

$$\hat{\tau}_{PCRD}^{BC,Linear} = \hat{\hat{\tau}}_Y = \hat{\tau}_{PCRD} - \left[(\hat{\hat{\tau}}_{\mathbf{Z},+})'\hat{\hat{\gamma}}_+ - (\hat{\hat{\tau}}_{\mathbf{Z},-})'\hat{\hat{\gamma}}_-\right] \xrightarrow{\mathbb{P}} (\tau_{PCRD} + b_{PCRD}) - b_{PCRD} = \tau_{PCRD}$$

Proof. See Appendix B.6. ■

Theorem 3 states that whenever 1) Assumption 2′ holds and 2) all accompanying characteristics that may induce the bias are observed, using the covariate-adjusted local polynomial estimator effectively identifies the target effect. However, to what extent are these results rooted in assuming a linear data generation process? I explore this question in the next section.

3.2 Bias correction in the general case

First, I would like to highlight the main takeaways from the linear case. On the one hand, by specifying an appropriate local model, local polynomial estimators can potentially eliminate the PCRD bias. On the other hand, this procedure yields an independent estimator of the bias $\hat{b}_{PCRD}^{Linear}$. These two findings motivate the following question: is it possible to modify the local model in such a way that the resulting local polynomial estimates replicate the properties exhibited in the linear case while imposing no functional form requirements?

Recall that regression functions are a general way of modeling the functional dependence between a politician's characteristics and the outcome. To extend these relationships to the district level, consider the following definitions.

$$m^{+}(u,\mathbf{Z}_d) = \lim_{\eta \downarrow 0} \mathbb{E}[Y_d|\mathbf{Z}_d, \Delta_d = u+\eta] \quad \text{and} \quad m^{-}(u,\mathbf{Z}_d) = \lim_{\eta \uparrow 0} \mathbb{E}[Y_d|\mathbf{Z}_d, \Delta_d = u+\eta]$$

In essence, $m^{+}(u,\mathbf{Z}_d)$ is the expected outcome level whenever the observed vote margin Δ_d is just above u and the winning candidate has attributes $\mathbf{Z}_d$. Thereby, in the particular case where $u=0$, the function $m^{+}(0,\mathbf{Z}_d(1))$ reflects the expected outcome level whenever candidate of type $X_{id}=1$ wins by a small margin and possesses attributes $\mathbf{Z}_d(1)$. Similarly, $m^{-}(0,\mathbf{Z}_d(0))$ describes a similar quantity, but in the case where the politician of type $X_{id}=0$ triumphs by a small margin and has characteristics $\mathbf{Z}_d(0)$. Nonetheless, the difference $m^{+}(0,\mathbf{Z}_d(1)) - m^{-}(0,\mathbf{Z}_d(0))$ will not be informative of a causal effect unless a comparability assumption, such as Assumption 1, holds. This condition is stated in Assumption 4.

Assumption 4. *The joint conditional expectation function of potential outcomes $Y_d(X_d)$, given by*

$$\mathbb{E}[Y_d(1)|\mathbf{Z}_d(1) = z, \Delta_d = u] \qquad \textit{and} \qquad \mathbb{E}[Y_d(0)|\mathbf{Z}_d(0) = z, \Delta_d = u],$$

are continuous at $u=0$ for any given z in the support of $\mathbf{Z}_d$.

Assumption 4 ensures that $m^{+}(0,\mathbf{Z}_d(1))$ and $m^{+}(0,\mathbf{Z}_d(0))$ are well-defined. Intuitively, this assumption postulates that candidates of either type that barely lose and barely win are comparable once all potential confounders are controlled. Therefore, this condition can be viewed as a relaxed version of Assumption 1 in that confounder imbalances at the threshold, such as compensating differentials, are possibly allowed.

Furthermore, Assumption 4 guarantees that *conditional* comparisons between politicians of opposing types account for proper causal effects. In other words, by fixing a characteristic vector z, then $m^{+}(0,z) - m^{-}(0,z) = \mathbb{E}[Y_d(1) - Y_d(0)|\mathbf{Z}_d = z, \Delta_d = 0]$, which represents the average treatment effect for electing a politician with $X_{id}=1$ and accompanying characteristics z by a small vote margin, over a candidate with $X_{id}=0$ with *the same attributes*. However, τ_{PCRD} is an *unconditional effect*, to the extent that it informs the effect of an average politician, regardless of its ancillary traits, that is involved in close races. Hence, it is necessary to prove that Assumption 4 is enough to identify the unconditional effect.

Theorem 4 (Frölich and Huber (2019a), Theorem 1)**.** *Consider Assumption 4 and regularity conditions RA-4, then τ_{PCRD} can be identified non-parametrically as*

$$\begin{aligned}
\tau_{PCRD} &= \mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0] \\
&= \mathbb{E}[m^{+}(0,\mathbf{Z}_d) - m^{-}(0,\mathbf{Z}_d)|\Delta_d = 0] \\
&= \int \left(m^{+}(0,z) - m^{-}(0,z)\right)\left(\frac{f^{+}_{\mathbf{Z}_d(0)|\Delta_d}(z|0) + f^{-}_{\mathbf{Z}_d(1)|\Delta_d}(z|0)}{2}\right) dz
\end{aligned}$$

where $f^{-}_{\mathbf{Z}_d(0)|\Delta_d}(\cdot|0)$ and $f^{+}_{\mathbf{Z}_d(1)|\Delta_d}(\cdot|0)$ are, respectively, the probability density functions of the random vectors $\mathbf{Z}_d(0)|\Delta_d = 0$ and $\mathbf{Z}_d(1)|\Delta_d = 0$[11].

Proof. See Appendix B.7. ∎

Theorem 4 attests that τ_{PCRD} is identified under Assumption 4. Furthermore, it offers a way of calculating τ_{PCRD} out of conditional comparisons, even when there is no information about the underlying data generation process. Essentially, the procedure entails averaging conditional average treatment effects using as weights the likelihood that each trait is observed in a tight electoral contest. However, since the potential traits can either originate from $\mathbf{Z}_d(1)$ or $\mathbf{Z}_d(0)$ in a close race, the probability of observing a particular attribute vector requires averaging two different distributions.

Now, is it possible to leverage the non-parametric identification of the causal effect to specify an appropriate local polynomial model? Consider a simple extension to model (3.2) given by

$$\dot{Y}_{id} = \dot{\alpha}_Y + X_d\dot{\tau}_Y + \Delta_{id}\dot{\beta}_{Y-} + (X_d \times \Delta_d)\dot{\beta}_{Y+} + X_d(\mathbf{Z}_{id} - \mathbf{c})'\dot{\gamma}_+ + (1 - X_d)(\mathbf{Z}_{id} - \mathbf{c})'\dot{\gamma}_- \tag{3.3}$$

where $\mathbf{c} \in \mathbb{R}^m$ is some fixed deterministic vector. By choosing this local model, we are intuitively asserting that

$$m^+(0, \mathbf{Z}_d) \approx \dot{\alpha}_Y + \dot{\tau}_Y + (\mathbf{Z}_d - \mathbf{c})'\dot{\gamma}_+$$
$$m^-(0, \mathbf{Z}_d) \approx \dot{\alpha}_Y + (\mathbf{Z}_d - \mathbf{c})'\dot{\gamma}_-$$

From which follows

$$m^+(0, \mathbf{Z}_d) - m^-(0, \mathbf{Z}_d) \approx \dot{\tau}_Y + (\mathbf{Z}_d - \mathbf{c})'(\dot{\gamma}_+ - \dot{\gamma}_-)$$

Thus, by taking expectations at both sides of the expression, Theorem 4 hints that

$$\tau_{PCRD} = \mathbb{E}[m^+(0, \mathbf{Z}_d) - m^-(0, \mathbf{Z}_d)|\Delta_d = 0] \approx \dot{\tau}_Y + (\mathbb{E}[\mathbf{Z}_d|\Delta_d = 0] - \mathbf{c})'(\dot{\gamma}_+ - \dot{\gamma}_-)$$

This expression implies that the covariate-adjusted local model will approximate the desired effect ($\dot{\tau}_Y \approx \tau_{PCRD}$) so long as *i*) $\dot{\gamma}_+ = \dot{\gamma}_-$, of which the "linear case" is a particular instance; and/or, *ii*) $\mathbf{c} = \mathbb{E}[\mathbf{Z}_d|\Delta_d = 0]$. On the one hand, *i*) is out of our control since its fulfillment depends on the (unknown) underlying data generation process. Condition *ii*), on the other hand, motivates a simple systematic method to remove bias: estimating model (3.3), replacing $\mathbf{c}$ with a consistent estimator of $\mathbb{E}[\mathbf{Z}_d|\Delta_d = 0]$. Such an estimator is given by[12]

$$\bar{\mathbf{Z}} = (\hat{\hat{\tau}}_{\mathbf{Z},+} + \hat{\hat{\tau}}_{\mathbf{Z},-})/2,$$

[11] In the case covariates are discrete, the result still holds by switching the corresponding integrals to sums and probability density functions to probability mass functions.

[12] A similar argument to Theorem 4's proof guarantees that

$$\bar{\mathbf{Z}} \xrightarrow{\mathbb{P}} \frac{\mu_{\mathbf{Z}+}(0) + \mu_{\mathbf{Z}-}(0)}{2} = \mathbb{E}[\mathbf{Z}_d|\Delta_d = 0]$$

through which we can specify the "recentered" covariate-adjusted local linear model as

$$\ddot{Y}_{id} = \ddot{\alpha}_Y + X_d \ddot{\tau}_Y + \Delta_{id} \ddot{\beta}_{Y-} + (X_d \times \Delta_d) \ddot{\beta}_{Y+} + X_d (\mathbf{Z}_{id} - \bar{\mathbf{Z}})' \ddot{\gamma}_+ + (1 - X_d)(\mathbf{Z}_{id} - \bar{\mathbf{Z}})' \ddot{\gamma}_- \tag{3.4}$$

Model (3.4) is essentially the same as model (3.2), except that observed confounders are demeaned using $\bar{\mathbf{Z}}$. As a result, rather than controlling for the politician's characteristics, the model considers how distinct these traits are from those of the typical politician running in close elections. Referring to the linear case, this is intuitively the correct specification because the bias is induced by relative endowments of characteristics rather than their absolute levels. Thus, flexibly controlling for "recentered" confounders may restore the identification of the unconditional causal effect. With minor changes, this idea has also been used to address similar issues in other causal identification methodologies, namely, in randomized controlled trials (Imbens & Wooldridge, 2009, Section 5.3), some instrumental variable designs (Borusyak & Hull, 2020), differences-in-differences (Wooldridge, 2021), and experiments featuring multiple treatments (Goldsmith-Pinkham, Hull, & Kolesár, 2022).

Consequently, the major result of this section is to demonstrate that the local polynomial estimator resulting from using model (3.4) can consistently estimate τ_{PCRD} regardless of the underlying regression function. I formalize these ideas in the following theorem.

Theorem 5. *Consider Assumption 4 and regularity conditions RA-1, RA-2, RA-3, RA-4 and RA-5 hold, then*

a)

$$\hat{\tau}^{BC}_{PCRD} = \hat{\ddot{\tau}}_Y = \hat{\tau}_{PCRD} - \frac{1}{2}\left[(\hat{\ddot{\tau}}_{\mathbf{Z},+} - \hat{\ddot{\tau}}_{\mathbf{Z},-})'(\hat{\ddot{\gamma}}_+ + \hat{\ddot{\gamma}}_-)\right]$$

b)

$$\hat{\tau}^{BC}_{PCRD} \xrightarrow{\mathbb{P}} \tau_{PCRD}$$

c)

$$\hat{b}_{PCRD} = \frac{1}{2}\left[(\hat{\ddot{\tau}}_{\mathbf{Z},+} - \hat{\ddot{\tau}}_{\mathbf{Z},-})'(\hat{\ddot{\gamma}}_+ + \hat{\ddot{\gamma}}_-)\right] \xrightarrow{\mathbb{P}} b_{PCRD}$$

Proof. See Appendix B.8.

■

Overall, Theorem 5 shows that the general unbiased estimator $\hat{\tau}^{BC}_{PCRD} = \hat{\ddot{\tau}}_Y$ has the same desirable qualities as its linear counterpart $\hat{\tau}^{BC,Linear}_{PCRD} = \hat{\dot{\tau}}_Y$, with the added benefit of not relying on a particular functional form of the regression function. Indeed, part $a)$ is the analog of Theorem 1, where covariate adjustment using model (3.4) introduces an additional term to the original PCRD estimator $\hat{\tau}_{PCRD}$. Next, in part $b)$, I prove that this extra bias term asymptotically removes PCRD estimation bias, thus mimicking the result of the linear case. Finally, in part $c)$, I conclude that the procedure yields a natural estimator for the bias, which

where $\mu_{\mathbf{Z}-}(0) = \mathbb{E}[\mathbf{Z}_d(0)|\Delta_d = 0]$ and $\mu_{\mathbf{Z}+}(0) = \mathbb{E}[\mathbf{Z}_d(1)|\Delta_d = 0]$. While any consistent estimator of $\mathbb{E}[\mathbf{Z}_d|\Delta_d = 0]$ will ensure overall consistency, using $\bar{\mathbf{Z}}$ significantly simplifies statistical inference and bias decompositions analysis, as well as leaves simple expressions when investigating bias sensitivity to omitted variables that will be introduced later. As a result, determining whether this is the best option for the estimator's asymptotic properties is beyond the scope of this paper.

can be used to study several decompositions of the bias term and do sensitivity analysis. The discussions in the next sections will revolve around these two directions.

Furthermore, for usual bandwidth choices, the covariate-adjusted estimator $\hat{\tau}^{BC}_{PCRD} = \hat{\hat{\tau}}_Y$ and the bias estimator $\hat{b}_{PCRD}$ are also asymptotically normal. To ease exposition, I leave optimal bandwidth choice and statistical inference details to Appendix B.9. Moreover, the resulting estimator is often more efficient than $\hat{\tau}_{PCRD}$. For instance, Calonico et al. (2019a, p. 448) prove that variance reduction is guaranteed whenever the influence of the additional covariates on the outcome near the cutoff is roughly the same for both control and treatment units ($\hat{\hat{\gamma}}_+ - \hat{\hat{\gamma}}_- \xrightarrow{\mathbb{P}} 0$). This indicates that $\hat{\tau}^{BC}_{PCRD}$ is potentially an overall better point estimate, in an MSE sense, than $\hat{\tau}_{PCRD}$ provided that it can simultaneously reduce bias and variance. Finally, the resulting estimator exhibits the same convergence rates as univariate nonparametric regressions ($n^{-2/5}$), which means its performance does not depend on the number or nature of the covariates (*i.e.* no curse of dimensionality). In Appendix A, I test some of these properties using Monte Carlo simulations.

Another significant implication of this result is that controlling for all relevant confounders can eliminate both confounding and post-treatment bias. While correlated characteristics are typically addressed by adjusting for omitted covariates, post-treatment bias is not usually considered to be resolved through covariate adjustment. Notably, my estimator can recover the actual causal effect despite the presence of both sources of bias. However, the ability to correct estimation bias remains dependent on observing all relevant confounders, a condition seldom or never met in practice. Even in the face of this constraint on its empirical application, constructing an estimate that solves the problem under specific circumstances is important because it allows benchmarking of bias in real-world scenarios. Section 4 will discuss this idea in depth.

3.3 Bias decomposition

In this section, I discuss why the bias term is an economically meaningful quantity and how Theorem 5 can help us understand the forces that drive it.

Recall that the traditional PCRD estimator $\hat{\tau}_{PCRD}$ reflects the outcome difference resulting from contrasting locations with and without an elected politician with characteristic X. Although this comparison does not fulfill the aim of PCRD designs of detecting the effect that having characteristic X has on the result, this does not mean that this quantity is not in itself economically meaningful. In fact, studying $\hat{\tau}_{PCRD}$ can still result very insightful for formulating public policy. To illustrate this point, I will formulate an example.

Returning to the question of whether women are more effective in positions of authority than men, let's assume that a specific study finds that *locations* in which women were elected had greater success in reducing domestic violence than those which appointed men ($\hat{\tau}_{PCRD} < 0$, when Y_d is the number of cases). Although the existence of a PCRD bias is likely, this finding informs a plain fact: places where women were elected, for whatever reason, had lower domestic abuse than those where men won the elections. As previously discussed, the part that remains obscure is, however, *why* are locations with elected women able to decrease this indicator. Is it because being a woman allows them to attain this reduction in domestic violence (τ_{PCRD})? Or is it because women in close races have distinctive characteristics that enable them to perform better in office (b_{PCRD})? Indeed, this partition is evident from the PCRD bias formula.

$$\hat{\tau}_{PCRD} \xrightarrow{\mathbb{P}} \underbrace{\tau_{PCRD}}_{\substack{\text{The part explained} \\ \text{by gender}}} + \underbrace{b_{PCRD}}_{\substack{\text{The part explained by other characteristics} \\ \text{correlated with gender in close races}}}$$

From the public policy standpoint, even when gender alone is not driving the differences, it would still seem desirable to know what explains the positive outcome in these locations. I will show that this knowledge can be acquired by decomposing the bias. Imagine a researcher uses the bias-corrected statistic in the described scenario and finds that $\hat{\tau}^{BC}_{PCRD} \approx 0$. Hence, it must be the case that other characteristics are driving the negative effect, leading to several relevant follow-up questions. For instance, which characteristics drive the negative effect?

Referring to the mnemonic presented in Section 3.1, the bias may originate from unbalanced characteristics between candidates in a close race (imbalance) or from differences in how these traits affect the outcome (impact). One possibility for the observed discrepancies is that women, on average, possess different relevant attributes that account for better performance in violence rates. Alternatively, these attributes may have a larger influence on the outcome (a larger marginal effect), leading to different results even when traits are the same. Understanding the source of the bias is critical for determining which characteristics of politicians are important in reducing domestic violence and hypothesizing what might happen in scenarios other than close races.

Suppose via a decomposition, we identify that the improvement in domestic violence indicators is attributable mainly to education levels. This is expected since women must compensate with other favorable characteristics, such as being more educated, to compete in close races. Now, imagine that an additional year of education is just as important in influencing the outcome for men as it is for women. Endowment differences would then need to explain the positive surplus: women outperform men in close races because they are more educated. As a result, whenever men and women with comparable education levels compete, no net effect on abuse rates can be anticipated because the effect that should prevail in this situation is the direct gender component, which is null ($\tau_{PCRD} \approx 0$). Alternatively, the reason why women outperform men in close races may be primarily because their attained education levels matter more for outcomes than men. In this circumstance, *gender does have a role*, to the extent that it reinforces the schooling effect. Considering this information, we can still anticipate a favorable effect if men and women compete with comparable levels of education based on this complementarity.

I will now show how these comparisons can be constructed from Theorem 5.

Decomposition by individual characteristics: The linearity of the model provides a way to identify separately the contribution of each characteristic. Namely, observe that

$$\hat{b}_{PCRD} = \sum_{l=1}^{m} \left[\frac{(\hat{\hat{\gamma}}_{-})_l + (\hat{\hat{\gamma}}_{+})_l}{2} (\hat{\hat{\tau}}_{X_d Z_{dl}} - \hat{\hat{\tau}}_{(1-X_d)Z_{dl}}) \right] = \sum_{l=1}^{m} \hat{b}_l$$

Hence, $\hat{b}_l$ is the part of the bias explained by characteristics Z_{idl}. Intuitively, as in the linear case, the amount of bias that each confounder contributes is "its impact times its imbalance". On the one hand, $\hat{\hat{\tau}}_{X_d Z_{dl}} - \hat{\hat{\tau}}_{(1-X_d)Z_{dl}}$ measures how dissimilar politicians of the type $X_{id} = 1$ and $X_{id} = 0$ involved in a close elections are. On the other hand, $(\hat{\hat{\gamma}}_{-})_l/2 + (\hat{\hat{\gamma}}_{+})_l/2$ describes the extent to which this imbalance impacts the outcome on average.

A "Kitagawa-Oaxaca-Blinder" decomposition: The Kitagawa–Oaxaca–Blinder decomposition (Blinder, 1973; Kitagawa, 1955; Oaxaca, 1973) is used to determine whether an observed difference between two groups is attributable to disparities in factor endowments, the differential impact of these factors on the outcome, or a combination. To derive a similar decomposition for the bias, observe that $\hat{\hat{\gamma}}_{+} = \hat{\hat{\gamma}}_{-} + (\hat{\hat{\gamma}}_{+} - \hat{\hat{\gamma}}_{-})$. Then, it follows that

$$\hat{b}_{PCRD} = \underbrace{\hat{\hat{\gamma}}_{-}' \left(\hat{\hat{\tau}}_{\mathbf{Z},+} - \hat{\hat{\tau}}_{\mathbf{Z},-}\right)}_{\text{Due to differences in endowments}} + \underbrace{\frac{1}{2} \left(\hat{\hat{\tau}}_{\mathbf{Z},+} - \hat{\hat{\tau}}_{\mathbf{Z},-}\right)' (\hat{\hat{\gamma}}_{+} - \hat{\hat{\gamma}}_{-})}_{\substack{\text{Due to differences in how} \\ \text{each type's attributes} \\ \text{impact the outcome}}}$$

Hypothesizing potential effects outside of close elections can be a particular benefit of doing this decomposition. For instance, imagine that differences in endowments primarily cause bias. Then, we could anticipate a difference in the outcome of about $\hat{\tau}^{BC}_{PCRD}$ in elections where candidates of opposing types are very similar. On the contrary, whenever bias is greatly explained by differences in the impact of these attributes, we cannot guarantee that $\hat{\tau}^{BC}_{PCRD}$ will be a good approximation of the net effect of the outcome even when candidates have comparable traits[13].

The "mixed" decomposition: The mixed decomposition combines the preceding partitions. To see how, remark that

$$\begin{aligned}
\hat{b}_{PCRD} &= \hat{\hat{\gamma}}_{-}' \left(\hat{\hat{\tau}}_{\mathbf{Z},+} - \hat{\hat{\tau}}_{\mathbf{Z},-}\right) + \left(\hat{\hat{\tau}}_{\mathbf{Z},+} - \hat{\hat{\tau}}_{\mathbf{Z},-}\right)' (\hat{\hat{\gamma}}_{+} - \hat{\hat{\gamma}}_{-})/2 \\
&= \sum_{l=1}^{m} (\hat{\hat{\tau}}_{X_d Z_{dl}} - \hat{\hat{\tau}}_{(1-X_d)Z_{dl}})(\hat{\hat{\gamma}}_{-})_l + ((\hat{\hat{\gamma}}_{+})_l - (\hat{\hat{\gamma}}_{-})_l)(\hat{\hat{\tau}}_{X_d Z_{dl}} - \hat{\hat{\tau}}_{(1-X_d)Z_{dl}})/2 \\
&= \sum_{l=1}^{m} \hat{E}_l + \hat{I}_l
\end{aligned}$$

[13]As noted by Fortin, Lemieux, and Firpo (2011), there can be some interpretational challenges for the KOB decomposition whenever there are categorical variables (with more than two categories). First, since categorical variables lack a natural zero, the reference point must be determined arbitrarily (*i.e.* which group to omit for estimation). The second issue is that omitting one category to determine the coefficients of the other categories renders it impossible to separate the part of the decomposition that can be attributed to group membership (at which we aim) from the part attributed to variations in the coefficient of the base or omitted category.

As in the aggregate decomposition, $\hat{E}_l$ represents the portion of $\hat{b}_l$ that is attributable to differences in endowments in Z_{idl}. Likewise, $\hat{I}_l$ accounts for the part of $\hat{b}_l$ that is explained by candidates of type $X_{id} = 1$ being able to influence more or less attribute Z_{idl} than politicians with $X_{id} = 0$ in the outcome.

The "mixed" decomposition is particularly useful to understand why a politician of type $X_{id} = 1$ is more or less successful than a candidate with $X_{id} = 0$. To put this idea into context, think about a study that focuses on the influence of a politician's pre-office vocation on some downstream outcome. For example, Kirkland (2021) and Szakonyi (2021) inquire whether businesspeople perform better in public office. Allegedly, having prior experience in management and a result-oriented mindset constitute an ideal skill set to govern, thus making businesspeople more efficient. Therefore, when a politician comes from a business background, their prior work experience and education may be relatively more important to their public success than in other circumstances. In particular, this mechanism would be supported by a net effect partially explained by a comparative advantage of businessmen and businesswomen using their education levels and years of work experience (*i.e.*, a large $\hat{I}_l$).

Finally, all decomposition components have a known asymptotic distribution as described in Appendix B.9, rendering statistical inference for each separate component possible.

4 Sensitivity analysis: can we assess how credible are PCRD designs?

In Section 3, I demonstrated that a covariate-adjusted local polynomial estimator could eliminate the PCRD estimation bias whenever all confounder characteristics are observed. However, in most applications, not all relevant candidate traits will be available in the data. For one, obtaining high-quality individual-level data on politicians may be difficult because such information is either restricted or nonexistent, particularly in low- and middle-income countries[14]. For another, many important individual attributes influencing public policy, such as motivation (Canon, 1993) or preparation (Carson, Engstrom, & Roberts, 2007), are by definition difficult to observe or measure.

Moreover, the set of possible omitted confounders can be enlarged by conditioning on close races. This is because compensating differentials can generate artificial differences in traits where there is no visible difference in the overall population. For instance, continuing on the gender example, while men and women are equally intelligent in the aggregate, women winning by a close margin over men may be relatively smarter than their male counterparts. As a result, intelligence would be a confounding factor in a PCRD design despite not being one in an observational study comparing outcomes by gender.

Considering these difficulties, covariate adjustment is unlikely to eliminate estimation bias fully. As a result, in addition to controlling for selection on observables, it would be desirable to conduct a "sensitiv-

[14]There are notable exceptions to this restriction, such as fine-grained administrative records in Sweden (See Dal Bó, Finan, Folke, Persson, and Rickne (2017)).

ity analysis" to determine how vulnerable a result is to unobserved confounding and post-treatment bias. Broadly, these methodologies have two main purposes. First, they aim to describe the type of unobserved confounders that would significantly alter the conclusions about the estimated causal effect. Second, based on the research design, available data, and expert knowledge, they assess the likelihood that such problematic confounding can exist. Combining these two pieces of knowledge can strengthen the validity of the estimated effects and create plausible bounds for the causal effect.

This section proposes a sensitivity analysis framework tailored to PCRD designs. My proposal is primarily inspired by Cinelli and Hazlett (2020a), who developed an analogous framework for addressing omitted variable bias in linear regressions, and which I will extend using the results of Section 3.

4.1 How problematic can an unobserved confounder be?

Let $\mathbf{Z}_d$ be the observed characteristics of the elected politician in district d. Suppose there exists a single unobserved trait U_d, which is a potential confounder to estimating the effect of attribute X_d on Y_d. It follows from Theorem 5 that estimating the local model

$$\begin{aligned}\ddot{Y}_{id} = \ddot{\alpha}_{Y,Full} + X_d \ddot{\tau}_{Y,Full} + \Delta_{id}\ddot{\beta}_{Y-,Full} + (X_d \times \Delta_d)\ddot{\beta}_{Y+,Full} + X_d(\mathbf{Z}_{id} - \bar{\mathbf{Z}})'\ddot{\gamma}_{+,Full} \\ + (1-X_d)(\mathbf{Z}_{id} - \bar{\mathbf{Z}})'\ddot{\gamma}_{-,Full} + X_d(U_{id} - \bar{U})\ddot{\rho}_{+,Full} + (1-X_d)(U_{id} - \bar{U})\ddot{\rho}_{-,Full}\end{aligned} \tag{4.1}$$

would yield a consistent estimator of the target causal effect insofar as $\hat{\ddot{\tau}}_{Y,Full} \xrightarrow{\mathbb{P}} \tau_{PCRD}$. However, since U_d is unobserved, researchers are limited to estimating the restricted model

$$\begin{aligned}\ddot{Y}_{id} = \ddot{\alpha}_{Y,Res} + X_d \ddot{\tau}_{Y,Res} + \Delta_{id}\ddot{\beta}_{Y-,Res} + (X_d \times \Delta_d)\ddot{\beta}_{Y+,Res} \\ + X_d(\mathbf{Z}_{id} - \bar{\mathbf{Z}})'\ddot{\gamma}_{+,Res} + (1-X_d)(\mathbf{Z}_{id} - \bar{\mathbf{Z}})'\ddot{\gamma}_{-,Res}\end{aligned} \tag{4.2}$$

Following this perspective, to assess the extent to which omitting a confounder can misdirect estimation, it is sufficient to understand how $\hat{\ddot{\tau}}_{Y,Res}$ compares with the desired estimate $\hat{\ddot{\tau}}_{Y,Full}$. Therefore, the discrepancy of these two quantities, given by $\widehat{\text{Bias}} = \hat{\ddot{\tau}}_{Y,Full} - \hat{\ddot{\tau}}_{Y,Res}$, is an estimate of the "residual" estimation bias after controlling for a set of observable variables. Moreover, observe that

$$\begin{aligned}\widehat{\text{Bias}} &= \hat{\ddot{\tau}}_{Y,Full} - \hat{\ddot{\tau}}_{Y,Res} \\ &= (\hat{\ddot{\tau}}_{Y,Full,+} - \hat{\ddot{\tau}}_{Y,Full,-}) - (\hat{\ddot{\tau}}_{Y,Res,+} - \hat{\ddot{\tau}}_{Y,Res,-}) \\ &= \underbrace{(\hat{\ddot{\tau}}_{Y,Full,+} - \hat{\ddot{\tau}}_{Y,Res,+})}_{\substack{\text{How different models are} \\ \text{at the right of the threshold.}}} - \underbrace{(\hat{\ddot{\tau}}_{Y,Full,-} - \hat{\ddot{\tau}}_{Y,Res,-})}_{\substack{\text{How different models are} \\ \text{at the left of the threshold.}}}\end{aligned}$$

which means that the bias can be alternatively described in terms of how different the estimators derived from each model are on either side of the threshold. Next, the linear structure of the bias-corrected estimator enables an explicit formulation of $\widehat{\text{Bias}}$ in terms of the complete and the restricted model. In particular, Proposition B.1 of Appendix B.10 attests that:

$$\hat{\hat{\tau}}_{Y,Res,+} - \hat{\hat{\tau}}_{Y,Full,+} = \hat{\hat{\rho}}_{+,Full}\hat{\hat{\tau}}_{U_d-\bar{U},+} \qquad \hat{\hat{\tau}}_{Y,Res,-} - \hat{\hat{\tau}}_{Y,Full,-} = \hat{\hat{\rho}}_{-,Full}\hat{\hat{\tau}}_{U_d-\bar{U},-}$$

where $\hat{\hat{\tau}}_{U-\bar{U},+}$ and $\hat{\hat{\tau}}_{U-\bar{U},-}$ would be the estimates at the right and the left side of the threshold of using model (3.4) on $U_d - \bar{U}$ as the dependent variable. As in the case of linear regressions, the former result provides an "omitted variable bias" formula that enables us to assess how the omission of covariates we wished to have controlled for could affect our inferences. Furthermore, notice the difference between the estimators has a similar structure to the decompositions in Section 3.3, where the bias depends on the imbalance of the confounder ($\hat{\hat{\tau}}_{U_d-\bar{U},+} - \hat{\hat{\tau}}_{U_d-\bar{U},-}$) and its differential impact on the outcome ($\hat{\hat{\rho}}_{+,Full} - \hat{\hat{\rho}}_{-,Full}$):

$$\widehat{\text{Bias}} = \underbrace{\hat{\hat{\rho}}_{+,Full}(\hat{\hat{\tau}}_{U_d-\bar{U},+} - \hat{\hat{\tau}}_{U_d-\bar{U},-})}_{\substack{\text{Bias due to imbalance}\\ \text{on unboservables}}} + \underbrace{\hat{\hat{\tau}}_{U_d-\bar{U},-}(\hat{\hat{\rho}}_{+,Full} - \hat{\hat{\rho}}_{-,Full})}_{\substack{\text{Bias due to differences}\\ \text{on how unboservables}\\ \text{influence the outcome}}}$$

Following the Frisch–Waugh–Lovell theorem, each bias component can be intuitively interpreted in terms of covariances and variances of "partialled-out" dependent outcomes at either side of the threshold. For instance, $\hat{\hat{\rho}}_{+,Full}$ can be intuitively described as

$$\hat{\hat{\rho}}_{+,Full} \approx \frac{\widehat{\text{Cov}}(X_d Y_d^{\perp(\mathbf{Z}_d-\bar{\mathbf{Z}})}, X_d[U_d-\bar{U}]^{\perp(\mathbf{Z}_d-\bar{\mathbf{Z}})}|\Delta_d = 0)}{\widehat{\text{Var}}(X_d[U_d-\bar{U}]^{\perp(\mathbf{Z}_d-\bar{\mathbf{Z}})}|\Delta_d = 0)}$$

where $Y_d^{\perp(\mathbf{Z}_d-\bar{\mathbf{Z}})}$ and $[U_d-\bar{U}]^{\perp(\mathbf{Z}_d-\bar{\mathbf{Z}})}$ represent variables Y_d and $U_d - \bar{U}$ after removing the components linearly explained by $\mathbf{Z}_d - \bar{\mathbf{Z}}$. Similar representations are possible for $\hat{\hat{\rho}}_{-,Full}$, $\hat{\hat{\tau}}_{U_d-\bar{U},+}$ and $\hat{\hat{\tau}}_{U_d-\bar{U},-}$.

In essence, deriving an explicit "omitted variable bias" formula for $\widehat{\text{Bias}}$ is useful because it allows us to identify which types of unobserved confounders will likely cause the most bias. The previous results suggest that the degree to which the omission of a confounder can distort results is determined by two affinities: *i*) how much does the unobserved confounder differential $U_d - \bar{U}$ correlate with the politician's type X_d (these come from the $\hat{\hat{\tau}}_{U_d-\bar{U},\pm}$); and *ii*), how strongly $U_d - \bar{U}$ is associated to the outcome Y_d (these result from the $\hat{\hat{\rho}}_{\pm,Full}$). More precisely, these associations are described by four quantities[15]:

- $R^+_{X_d \sim U_d-\bar{U}|\mathbf{Z}_d-\bar{\mathbf{Z}}} :=$ The estimated residual correlation between X_d and $U_d(1) - \bar{U}$ after controlling for observed covariates $\mathbf{Z}_d(1) - \bar{\mathbf{Z}}$, when $\Delta_d = 0$.
- $R^+_{Y_d \sim U_d-\bar{U}|X_d,\mathbf{Z}_d-\bar{\mathbf{Z}}} :=$ The estimated residual correlation between Y_d and $U_d(1) - \bar{U}$ after controlling for observed covariates $\mathbf{Z}_d(1) - \bar{\mathbf{Z}}$ and X_d, when $\Delta_d = 0$.
- $R^-_{X_d \sim U_d-\bar{U}|\mathbf{Z}_d-\bar{\mathbf{Z}}} :=$ The estimated residual correlation between X_d and $U_d(0) - \bar{U}$ after controlling for observed covariates $\mathbf{Z}_d(0) - \bar{\mathbf{Z}}$, when $\Delta_d = 0$.

[15] For precise mathematical definitions, please refer to the Appendix B.11.

- $R^{-}_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d - \bar{\mathbf{Z}}}$:= The estimated residual correlation between Y_d and $U_d(0) - \bar{U}$ after controlling for observed covariates $\mathbf{Z}_d(0) - \bar{\mathbf{Z}}$ and X_d, when $\Delta_d = 0$.

In Appendix B.11, I show that these quantities completely characterize the bias to the extent that

$$\widehat{\text{Bias}} = \hat{\mathcal{C}}_{+} \times \left(\frac{R^{+}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}} \times R^{+}_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d - \bar{\mathbf{Z}}}}{\sqrt{1 - [R^{+}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}}]^2}} \right) + \hat{\mathcal{C}}_{-} \times \left(\frac{R^{-}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}} \times R^{-}_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d - \bar{\mathbf{Z}}}}{\sqrt{1 - [R^{-}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}}]^2}} \right)$$

where $\hat{\mathcal{C}}_{+}$ and $\hat{\mathcal{C}}_{-}$ are constants that can be estimated using only information from the restricted model[16]. Consequently, the complete incidence of unobserved confounders in the estimated effect is summarized by two quantities, known as "Bias Factors", which are embodied in the model as

$$\text{BF}^{+} = \frac{R^{+}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}} \times R^{+}_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d - \bar{\mathbf{Z}}}}{\sqrt{1 - [R^{+}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}}]^2}}; \quad \text{BF}^{-} = \frac{R^{-}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}} \times R^{-}_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d - \bar{\mathbf{Z}}}}{\sqrt{1 - [R^{-}_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d - \bar{\mathbf{Z}}}]^2}}$$

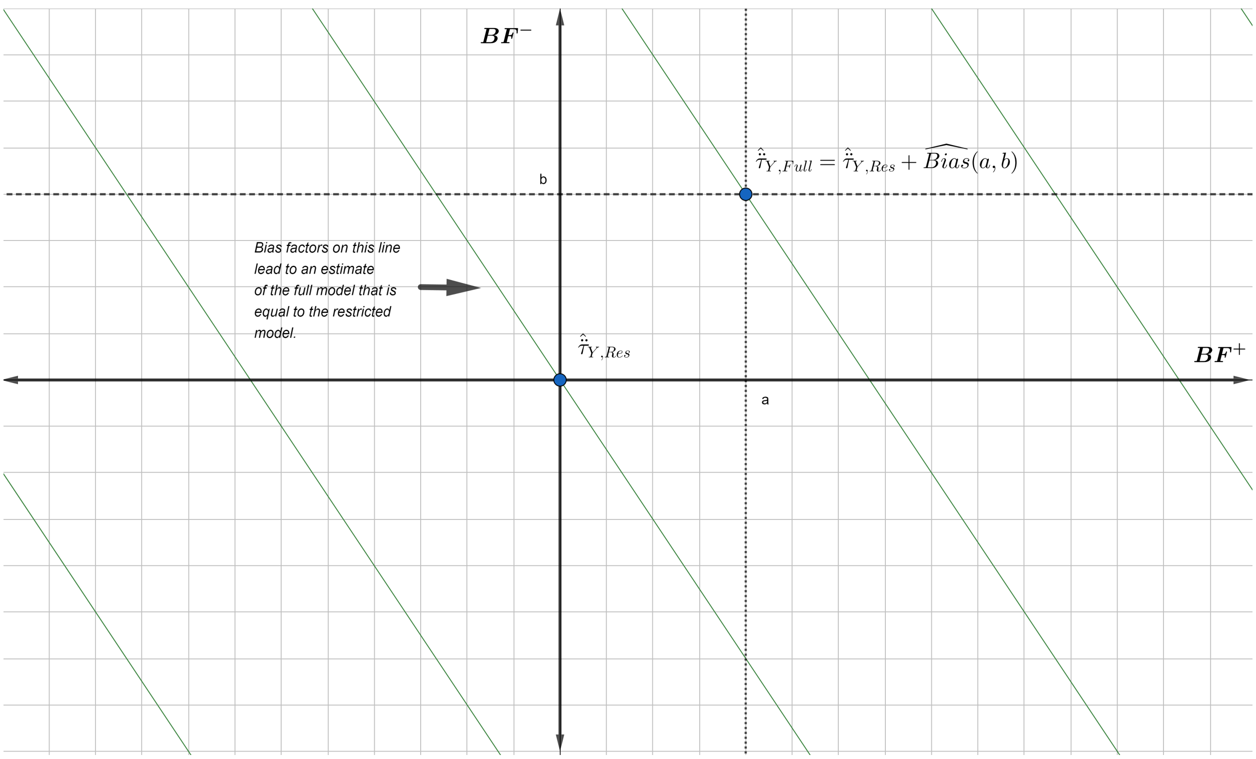
Following these results, measuring the amount of bias induced by a specific confounder's omission is equivalent to determining its two bias factors. Moreover, these quantities hold an intuitive interpretation: they are an aggregate measure of how relevant differences in the omitted characteristic are, *rather than* X_d, at determining outcome levels at the right (BF^{+}) or at the left (BF^{-}) of the threshold. Finally, notice that these quantities intuitively combine the correlation strength between the unobservable and the trait (*i.e.*, the imbalance part) and the correlation strength between the unobservable and the outcome (*i.e.*, the impact part) to account simultaneously for both sources of bias.

Likewise, the omitted attribute bias can be seen as a function of these two variables, *i.e.*, $\widehat{\text{Bias}} = \widehat{\text{Bias}}(\text{BF}^{+}, \text{BF}^{-})$. This formulation admits a convenient graphical depiction of the bias, known as *sensitivity contour plots* (Cinelli & Hazlett, 2020a). Essentially, these diagrams leverage the fact that each pair of hypothetical bias factors $(\text{BF}^{+}, \text{BF}^{-})$ corresponds to a particular level of bias. Hence, given an initial, albeit incorrect, estimate of the causal effect $\hat{\hat{\tau}}_{Y,Res}$, it is possible to reproduce the estimate that we would have obtained using the complete model via the equation $\hat{\hat{\tau}}_{Y,Full} = \hat{\hat{\tau}}_{Y,Res} + \widehat{\text{Bias}}(\text{BF}^{+}, \text{BF}^{-})$. Thereby, using a contour plot[17], it is possible to illustrate the particular values of $\hat{\hat{\tau}}_{Y,Full}$ associated to each point in the $\text{BF}^{+} - \text{BF}^{-}$ plane. In particular, it is possible to determine the set of points for which $\hat{\hat{\tau}}_{Y,Full} = 0$ or for which the estimated sign flips, which reflect circumstances in which unobservable confounding dramatically modifies the study conclusions. An illustration of this is available in Figure 2.

[16] These constants result from transforming sample covariances into sample correlations and from manipulating partial correlations.

[17] A contour plot, or level plot, is a graphical technique for representing a 3-dimensional surface $z = \phi(x, y)$ in 2-dimensional space by plotting constant z slices over the $x - y$ plane. In other words, it shows the set of points (x, y) that yield the same value of $z = \phi(x, y)$, for several z in the range of ϕ.

FIGURE 2: *An example contour plot illustrating bias factor incidence on an observed estimate.*



Notes: Hypothetical bias contour plot for positive $\hat{\mathcal{C}}_+$ and $\hat{\mathcal{C}}_-$. The $X-$axis reflect values of BF^+, while the $Y-$axis values of BF^-. Green lines represent level curves of the function $\widehat{Bias}(BF^+, BF^-)$. Every point in the plane represents the estimate of the "Full" model corresponding to a tuple of bias factors. For instance, I indicate the value of the "Full" model when bias factors are some arbitrary pair (a, b). When there is no bias due to unobservables—represented by the point $(0, 0)$—the estimate resulting from the "Full" model ($\hat{\ddot{\tau}}_{Y,Full}$) matches that observed in the "Restricted" model ($\hat{\ddot{\tau}}_{Y,Res}$).

4.2 Can we assess whether a problematic omitted confounder exists in a particular study?

While it is possible to quantify exactly how influential an unobserved trait must be for the results to be discredited, assessing whether such a confounder might exist in one's study remains a challenge. This is because assessing the maximum strength of confounders typically involves some subjective assertion, informal benchmark, or "back of the envelope" calculation. Therefore, many popular sensitivity analysis frameworks, such as Imbens (2003) and Oster (2019), suggest that the best way to assess the influence of omitted confounders is by benchmarking its possible strength using other observed covariates. In essence, the rationale for the method is that even if a researcher cannot directly assess the influence of an unobserved confounder on the bias, he or she may still be able to assess its *relative* strength, for example, by claiming that the omitted characteristic cannot possibly account for as much variation as some other observed trait.

For instance, let Z_{idl} be the benchmark characteristic[18] and denote by

$$\mathbf{Z}_{id}^{-l} = (Z_{id,1}, \cdots, Z_{id,l-1}, Z_{id,l+1}, \cdots, Z_{id,m})'$$

the observed trait vector excluding attribute Z_{idl}. Now let

$$\kappa_X^+ := \left(\frac{R^+_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}}{R^+_{X_d \sim Z_{dl} - \bar{Z}_l | \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}} \right)^2; \quad \kappa_Y^+ := \left(\frac{R^+_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}}{R^+_{Y_d \sim Z_{dl} - \bar{Z}_l | X_d, \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}} \right)^2$$

In words, κ_X^+ represents how much more of the variance of X_d does $U_d - \bar{U}$ explain relative to the benchmark difference $Z_{dl} - \bar{Z}_l$ (after controlling for the remaining characteristics)[19]. κ_Y^+ has a similar interpretation but instead highlights the share of the variance of the outcome variable Y_d at the right of the cutoff is explained by $U_d - \bar{U}$ with $Z_{dl} - \bar{Z}_l$, after controlling for the other observed traits. Alternatively, κ_X^+ and κ_Y^+ account for how much more is $U_d - \bar{U}$ a better predictor of X_d and $X_d Y_d$ than $Z_{dl} - \bar{Z}_l$, respectively.

Likewise, it is possible to define analogous quantities for the left side of the threshold as

$$\kappa_X^- := \left(\frac{R^-_{X_d \sim U_d - \bar{U} | \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}}{R^-_{X_d \sim Z_{dl} - \bar{Z}_l | \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}} \right)^2; \quad \kappa_Y^- := \left(\frac{R^-_{Y_d \sim U_d - \bar{U} | X_d, \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}}{R^-_{Y_d \sim Z_{dl} - \bar{Z}_l | X_d, \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}} \right)^2$$

Without loss of generality, consider U_d to be the part of U_d not linearly explained by $\mathbf{Z}_d$ (so that U_d and $\mathbf{Z}_d$ are uncorrelated). Following this characterization, I show in Appendix B.12 that bounding the

[18]A possible choice of this trait can be based on the bias decompositions of Section 3. For instance, one could choose the most influential attribute as the benchmark.

[19]This assertion follows from the same logic of $R-$squared statistics in linear regressions. Given a linear regression model $Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$, let $\hat{Y}_i$ and $\hat{\varepsilon}_i$ be the model's predicted outcome and residuals respectively. Then

$$R^2_{Y\sim D} = \underbrace{\frac{\text{Var}(\hat{Y}_i)}{\text{Var}(Y_i)}}_{\text{Share of the variance of } Y_i \text{ explained by } D_i.} = 1 - \frac{\text{Var}(\hat{\varepsilon}_i)}{\text{Var}(Y_i)} = \text{Corr}(\hat{Y}_i, Y_i)^2 = \underbrace{\text{Corr}(Y_i, D_i)^2}_{\text{Square of the correlation between the dependent and the independent variables.}}$$

When additional controls exist, substitute Y_i and D_i with their "partialed-out" versions.

bias factors in terms of these quantities is possible. More precisely, there exist functions $\overline{\text{BF}}^+(\kappa_X, \kappa_Y)$ and $\overline{\text{BF}}^-(\kappa_X, \kappa_Y)$ such that

$$|\text{BF}^+| \leq \overline{\text{BF}}^+(\kappa_X^+, \kappa_Y^+); \quad |\text{BF}^-| \leq \overline{\text{BF}}^-(\kappa_X^-, \kappa_Y^-)$$

Furthermore, these constraints are tight, meaning they are the smallest possible upper bounds for these numbers. Finally, apart from redirecting the focus from absolute to relative terms, this approach also makes it possible to use the available data to infer the maximum conceivable strength of any potential confounders. For instance, if we estimated that $[R^+_{X_d \sim Z_{dl} - \bar{Z}_l | \mathbf{Z}_d^{-l} - \bar{Z}^{-l}}]^2 > 0.5$, this is, the differential of the benchmark attribute explained more than half of the residual variation of X_d just at the right of the threshold, then no other characteristic could possibly be as strong as this characteristic in this dimension. As a result, *data would disclose* that the only admissible values for κ_X^+ would be those between 0 and 1.

Following this result, it is also possible to cap the aggregate residual bias to the extent that

$$|\widehat{\text{Bias}}(\mathbf{h})| \leq |\hat{\mathcal{C}}_+| \, \overline{\text{BF}}^+(\kappa_X^+, \kappa_Y^+) + |\hat{\mathcal{C}}_-| \, \overline{\text{BF}}^-(\kappa_X^-, \kappa_Y^-) = \overline{\text{Bias}}(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$$

Yet, even when it is possible to reason separately about each component of $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$, Cinelli and Hazlett (2020a) propose the clearest analysis results from assuming that $\kappa_X^+ = \kappa_Y^+ = \kappa_X^- = \kappa_Y^- = \kappa$. Doing so reduces the problem to determining *how much point estimates could change if the unobserved confounder was at most κ times as strong as the benchmark characteristic.* Therefore, bounds for the desired effect can be constructed using the fact that

Yet, even when it is possible to reason separately about each component of $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$, the clearest analysis results from assuming that $\kappa_X^+ = \kappa_Y^+ = \kappa_X^- = \kappa_Y^- = \kappa$. Doing so reduces the problem to determining *how much point estimates could change if the unobserved confounder was at most κ times as strong as the benchmark characteristic.* Therefore, bounds for the desired effect can be constructed using the fact that

$$\hat{\hat{\tau}}_{Y,Res} - \overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa) \leq \hat{\hat{\tau}}_{Y,Full} \approx \tau_{PCRD} \leq \hat{\hat{\tau}}_{Y,Res} + \overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa)$$

In Appendix B.13, I demonstrate how the single unobserved confounder framework used in this section is general enough to cover the case of multiple omitted confounders. Furthermore, I extend the bounding methodology to allow for benchmarks that include multiple or all observed characteristics. Finally, in Appendix C.1, I explain how this procedure produces valid bounds for effect, regardless of the nature of $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$.

A potential complication for adopting this analysis is that, whereas bias depends on the relative strength between unobservables and the selected benchmark *in a close election*, practitioners may only have a sense of this at the population level. Because conditioning on tight elections can amplify disparities between politicians on unobservable dimensions—thereby increasing relative strength from that known in aggregate—these two numbers can dramatically fluctuate due to post-treatment bias. This poses problems for interpretation, as using the population's relative strength as a reference may underestimate the possible magnitude of the bias. Moreover, holding fixed other variables can further exacerbate the post-treatment bias effect in the unobserved dimension. However, as shown in Appendix C.2 relative strengths are not guaranteed to be inflated by conditioning; they can also be deflated depending on the choice of benchmarking category. These concerns can be alleviated by selecting benchmark variables that are also significantly influenced by conditioning. The intuition behind this result is that relative strength is a ratio. Thus, even when the numerator (the unobservable) is likely to increase due to post-treatment bias, choosing a denominator (the benchmark) that is also heavily influenced by conditioning can help balance out the effect, bringing the ratio closer to the population benchmark.

5 From theory to practice: performance of female mayors in the pandemic.

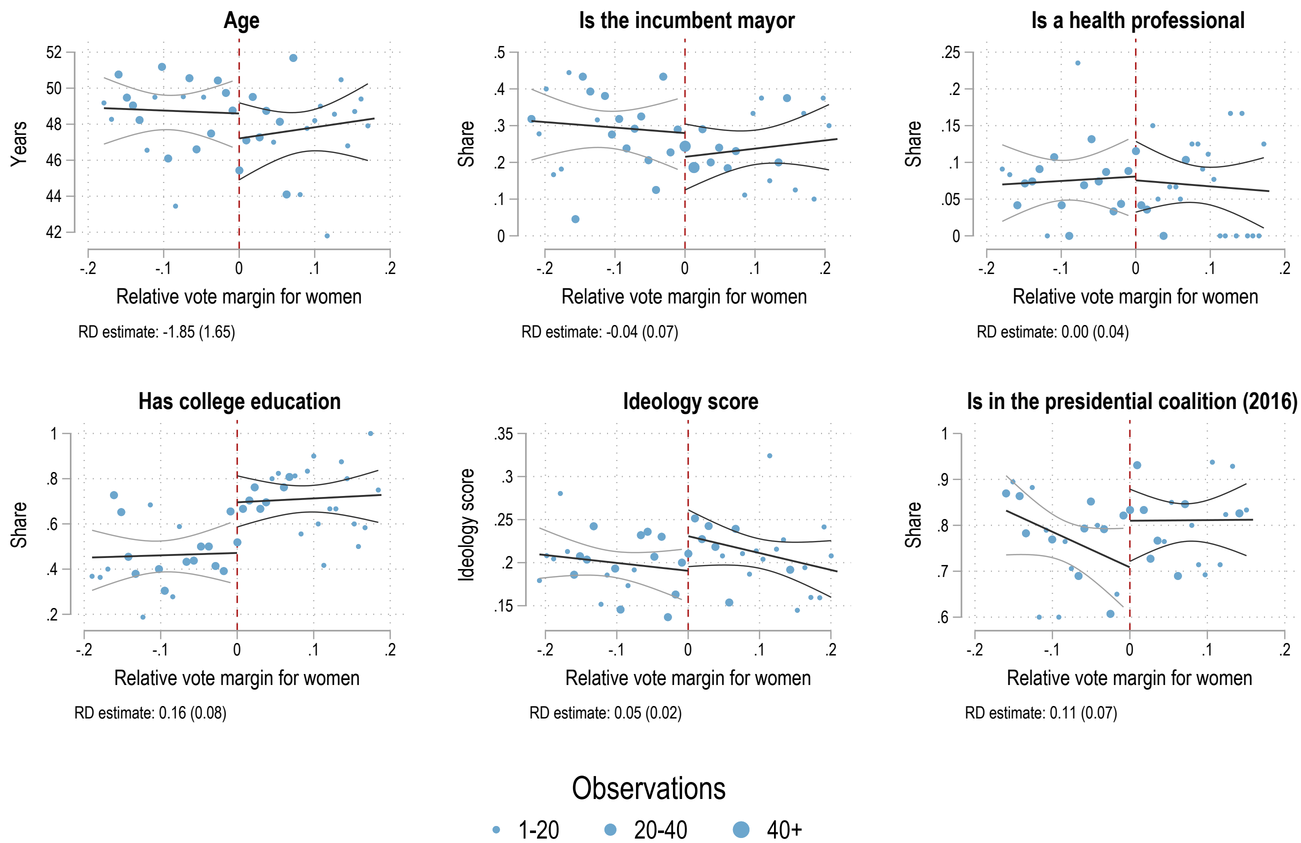
This section demonstrates the proposed econometric framework through a case study. Specifically, it revisits the article "*Under pressure: Women's leadership during the COVID-19 crisis*" by Bruce et al. (2022), published in the *Journal of Development Economics*. Specifically, I constructed a database using the raw sources listed by the authors and included additional politician characteristics from the *Tribunal Superior Eleitoral*, Brazil's Electoral authority. Reassuringly, my analysis using this dataset yielded the same qualitative results as the original paper.

This paper examines the impact of female mayors during the COVID-19 pandemic in Brazil. The authors' identification strategy is a PCRD design, leveraging close mayoral races between male and female candidates in 2016 as a potential source of quasi-random variation. Next, they investigate whether places with female leadership had lower rates of deaths and hospitalizations due to COVID-19 and other Severe Acute Respiratory Infections (SARI) than places with male leadership. The data used for estimation comes from a variety of sources including the SIVEP-Gripe system, Brazil's Electoral Court, and the data from Power and Rodrigues-Silveira (2019).

As discussed in this paper, this particular PCRD will likely confound the effect of compensating differentials and the effect attributable to gender. This observation becomes clearer when comparing the individual traits of candidates of different genders involved in close races. Figure 3 suggests that narrow female winners are more likely to be college-educated, are slightly younger, more right-wing oriented, and more likely to belong to the presidential coalition in 2016 than their male counterparts. Despite the presence and awareness of potential confounders, Bruce et al. (2022) still claim that "female leadership reduced deaths and hospitalizations per 100 thousand inhabitants. (p.1)".

Yet, to what extent can we be confident that differences in epidemiological outcomes are driven by gender, as claimed by the authors, rather than other characteristics of candidates in close races? A first step towards answering this question is correcting for potential bias due to unbalanced observable traits. Bruce et al. (2022) attempt this by using the estimator proposed by Calonico et al. (2019a) and conclude that their results are not due to "other observable mayoral characteristics such as education or political preferences

FIGURE 3: *Compensating differentials in Bruce et al. (2022)*



Notes: RD plots of several individual characteristics of elected politicians in Brazilian municipal governments in 2016 that competed in close races involving candidates of different genders. The $X-$axis accounts for the difference in votes between the most voted women and the most voted men in the contest. Bins were selected following the mimicking variance evenly-spaced method developed by Calonico et al. (2015) using spacings estimators. Local linear polynomial estimates and robust standard errors (in parenthesis) are reported following Calonico et al. (2014) using a triangular kernel and an MSE-optimal bandwidth. The ideology score in the lower center panel is measured from $[-1, 1]$, where a positive score indicates a right-wing position, whereas negative values represent left-wing political beliefs.

(p.1).". However, as shown in Section 3.1, this assertion overlooks that the local model used in Calonico et al. (2019a) would only be effective at correcting the bias for some specific regression functions. This is because the employed local model does not allow enough flexibility to be a complete non-parametric estimator of the local average treatment effect in the presence of compensating differentials.

Table 1 shows the covariate-adjusted estimators following the bias correction procedure proposed in Section 3.2. Notably, the original estimates (Panel A) and the politician-attribute corrected estimates (Panel B) are relatively similar in size and still statistically significant at the 0.05 level for all the epidemiological outcomes. Panel C emphasizes this observation, given that the independent estimate of the PCRD bias is small and not statistically significant at the reference level. Thus, no sizeable bias appears to arise from observable traits despite the significant differentials in some traits exhibited in Figure 3.

To further investigate the consequences of observable compensating differentials on the estimate, I conduct the decompositions introduced in Section 3.3 for COVID-19 deaths (Column 1 of Table 1). Figure 4 summarizes these results. While the estimated bias is close to zero and non-significant, as suggested in Table 1, the decompositions reveal that their components are not. On the one hand, the fact that narrow female winners are more ideologically right-wing oriented than males produces a sizeable (12.5% the size of the uncorrected coefficient) and statistically significant bias at the 0.05% level. However, the sign of the bias suggests that not accounting for ideological differences would underestimate the role of gender leadership, meaning that places where right-wing-leaning politicians were elected fared better on average than places with candidates on the opposite side of the political spectrum[20]. Moreover, the "mixed" decomposition reveals that the bias is due to differences in ideological positions rather than right/left-wing males influencing outcomes differently than right/left-wing female politicians.

On the other hand, other characteristics that exhibited differentials, such as education and belonging to the presidential coalition, indicate that the effect attributed to gender may be overestimated. For instance, considering that females involved in close elections are better educated would decrease the estimated effect of the uncorrected coefficient by 10%. However, this reduction is statistically non-significant. Additionally, the "mixed" decomposition reveals that while differences in education endowments explain a bigger share of the effect (15%), educated men seem to outperform women of similar education levels, making the aggregate bias smaller. Similar results hold when considering membership in the government coalition, whose incidence in the effect is positive and statistically significant at the 0.1 level. Thus, it is not that observed differentials do not alter results, but rather that the biases they generate cancel out. This demonstrates a major advantage of the method: it not only establishes whether the bias is zero in a particular application but also explains why, thus delivering meaningful implications even under zero bias.

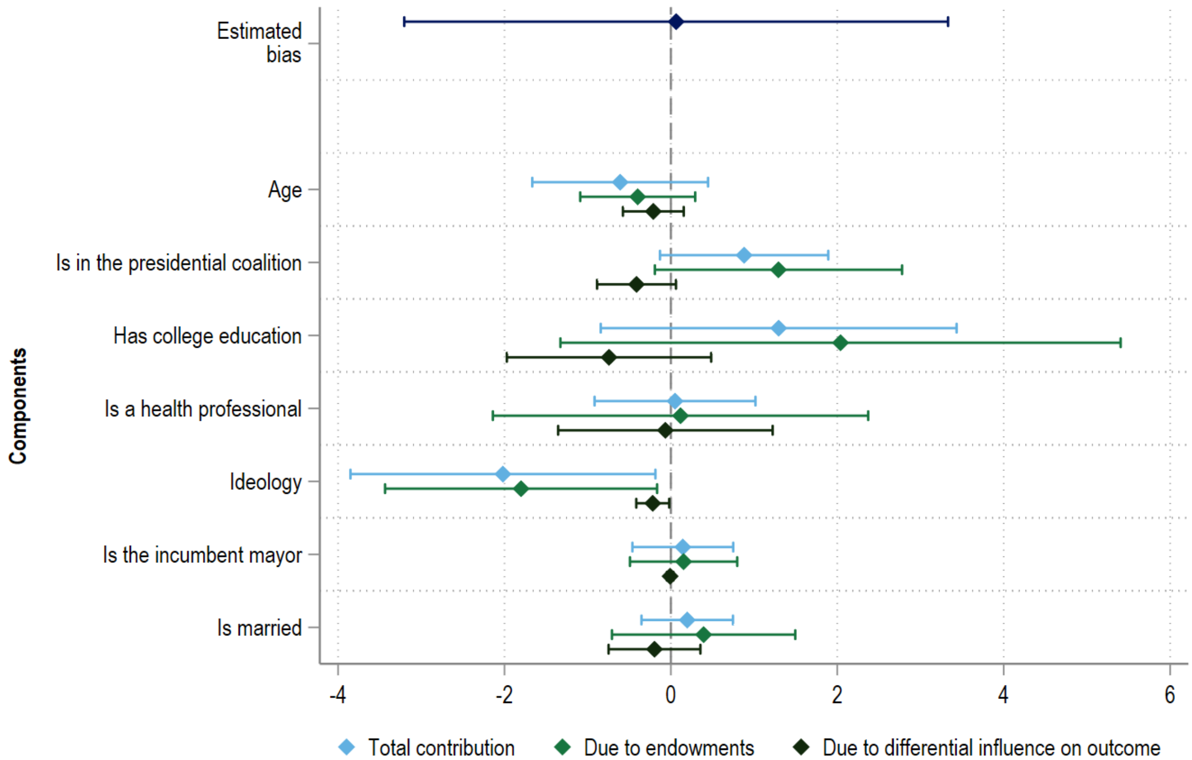
[20]It is important to note that the decompositions of the bias cannot be interpreted causally. The results indicate that better epidemiological indicators were observed in places where right-wing politicians won. However, this does not mean that ideology accounts for the positive results observed in these places. These quantities are merely correlations, and any policy implications should be gauged accordingly.

TABLE 1: *Traditional vs. covariate-adjusted estimators.*

	(1) **COVID-19 deaths per 100k pop.**	(2) **COVID-19 hosp. per 100k pop.**	(3) **SARI deaths per 100k pop.**	(4) **SARI hosp. per 100k pop.**
Panel A: Original estimates				
Female wins [1=Yes]	-16.224	-52.732	-22.648	-87.515
	(6.287)	(20.617)	(8.144)	(32.648)
	[-30.97,-6.32]	[-100.64,-19.82]	[-41.75,-9.82]	[-161.77,-33.79]
Bandwidth size	0.161	0.148	0.151	0.147
Observations	1238	1238	1238	1238
Observations in Bandwidth	780	731	740	723
Panel B: Politician-attribute corrected estimate.				
Female wins [1=Yes]	-16.809	-54.729	-23.236	-92.307
	(6.337)	(20.433)	(8.079)	(30.132)
	[-31.12,-6.28]	[-100.02,-19.93]	[-41.34,-9.67]	[-161.65,-43.53]
Bandwidth size	0.144	0.148	0.141	0.146
Observations	1238	1238	1238	1238
Observations in Bandwidth	709	732	690	722
Panel C: Estimated bias on observables.				
Female wins [1=Yes]	-0.248	-1.289	-0.737	-3.757
	(1.667)	(4.775)	(2.255)	(7.596)
	[-3.52, 3.02]	[-10.65, 8.07]	[-5.16, 3.68]	[-18.64,11.13]
Bandwidth size	0.144	0.148	0.141	0.146
Observations	1238	1238	1238	1238
Observations in Bandwidth	709	732	690	722

Notes: This table presents estimates for the impact of female victory on COVID-19 and SARI deaths and hospitalizations. Panel A shows Calonico et al. (2014) estimates under a linear model, triangular kernel, and MSE-optimal bandwidth choice. Panel B contains covariate corrections for differences in age, incumbency, marriage status, ideology, being a health professional, national government coalition membership, and college education. Panel C is the estimated difference between Panel A and Panel B. Robust standard errors in parenthesis and 95% confidence intervals in squared brackets. All specifications include state fixed effects.

FIGURE 4: *Bias decomposition COVID-19 deaths per hundred-thousand people.*

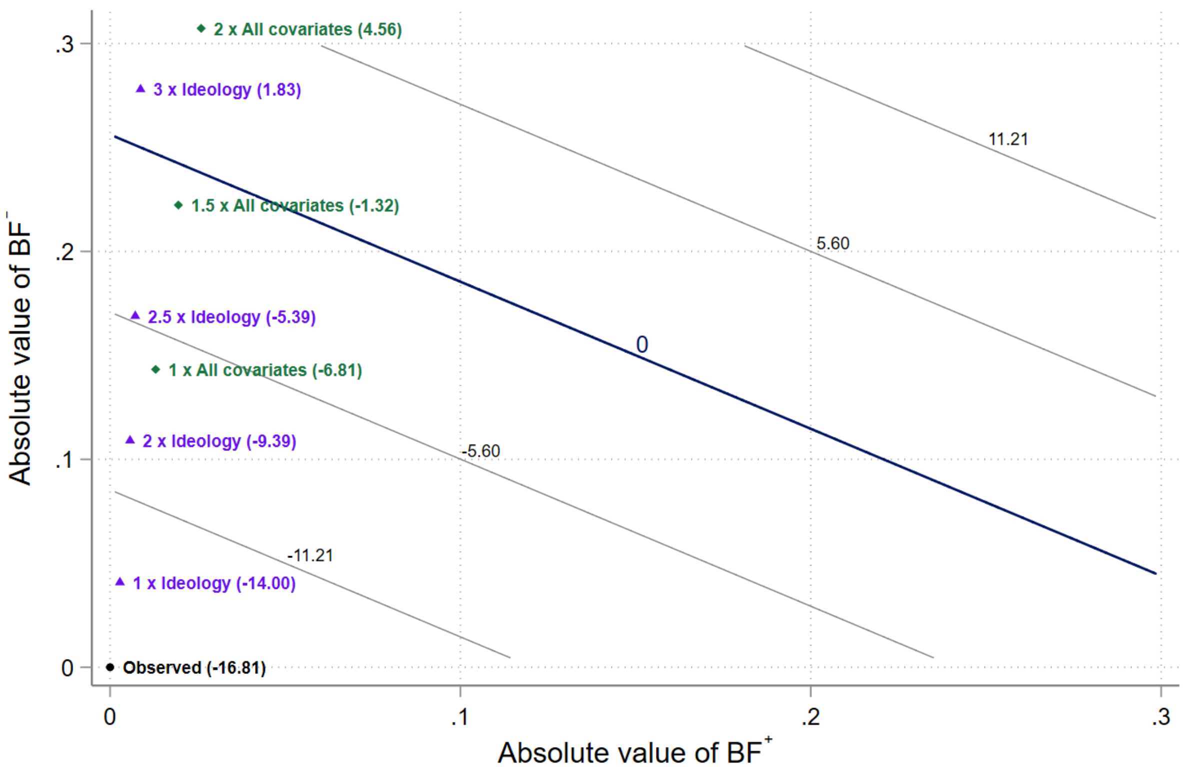


Notes: The dark blue coefficient shows the estimated bias on observables. Rows 2-7 show how individual traits contribute to this bias. Pink estimates show each trait's aggregate contribution to the bias. Green and brown estimates further break down this quantity (pink estimate) into the part attributable to endowment differences and the part ascribable to differences in the impact on the outcome of said characteristic, respectively. All point estimates include state-level fixed effects and are recentered. 95% confidence intervals.

Even when the analysis of observed differentials seems to support that gender differences drive the estimated effect, lurking in the background is the possibility of an unobserved confounder explaining the effect. To assess this prospect, I will use the results derived in Section 4 to test the possible effects that unobserved confounders using two formal benchmarks: ideology, which is the most influential observed confounder according to the decomposition, and also, the whole set of observed traits combined.

Figure 5 shows the contour plot of the estimated coefficients resulting from different bias factors combinations as proposed in Section 4. More precisely, it displays level curves for different estimate values and the combinations of bias factors that would result in zero estimated coefficients (highlighted in dark blue). I also plotted the maximum admissible coefficient estimates that would result from omitting a confounder that is k times as strong as the benchmarks. For instance, an unobserved confounder 2.5 times stronger than ideology, generating a positive bias, would not nullify the result (at most, it would result in an estimated effect of -5.39). However, a confounder three times stronger than ideology could result in a positive estimate, thus invalidating the authors' conclusions. Likewise, a confounder 1.5 stronger than all observed confounders could not disprove the authors' conclusions, but one two times stronger has the potential to do so.

FIGURE 5: *Sensitivity to unobservables: contour plot with formal benchmarks.*

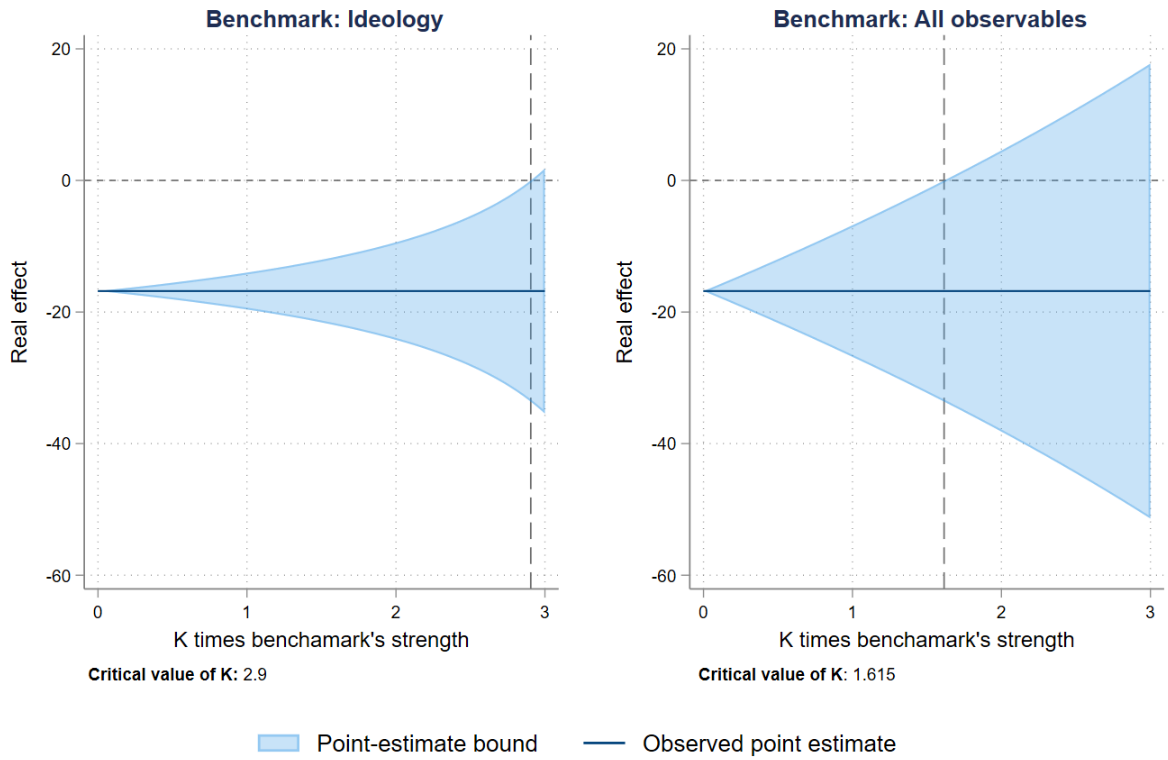


Notes: This graph shows a contour graph for the estimated effect of the estimated effect that would be retrieved if an unobserved confounder with bias factors $(\text{BF}^+, \text{BF}^-)$ was observed and corrected for. Each line represents a different level curve for the point estimates. The dark blue line indicates the combinations of bias factors that would lead the estimated effect to be zero. Blue triangles show the bias factors and corrected point estimates of a confounder $k-$times as strong as ideology. Green diamonds do the same for confounder $k-$times as strong as all observed politician traits.

A finer-grained analysis is provided in Figure 6, where I plot the implied bounds in the estimated effect for different levels of k, using both benchmarks. For instance, any unobserved confounder less than 2.9 times as strong as ideology would lead to the estimated effect remaining negative. Similarly, any omitted characteristic less than 1.6 times as strong as all the observed variables would have the same result.

I interpret these findings as evidence of the authors' claim being relatively robust to unobserved confounding. On the one hand, ideology is the most influential of all observed confounders, surpassing other traits spotlighted in the literature to be important for public policy, such as age (Alesina, Cassidy, & Troiano, 2019), education (Besley, Montalvo, & Reynal-Querol, 2011) and political experience (Jones, Saiegh, Spiller, & Tommasi, 2002). Thus, it seems unlikely a confounder **almost thrice** as strong as ideology exists, provided most of the potential suspects are already included in the model.

FIGURE 6: *Point-estimate bounds using formal benchmarks for various values of k.*



Notes: Point estimate bounds for different values of k, using as formal benchmarks ideology and all observed traits.

On the other hand, other sensitivity methodologies like Oster (2019) and Altonji et al. (2005) suggest that an estimate is robust to unobservables if an omitted confounder **as strong as the complete set of controls** cannot drastically change the article's conclusions. Hence, according to this criterion, these results would be relatively robust to omitted variables.

However, assessing the sensitivity of estimates to unobservables should always be gauged in context. For instance, the rationale behind the previous empirical rule is that researchers typically focus their data collection efforts (or their choice of regression controls) on the variable they believe *ex-ante* are the most important to adjust for. Consequently, it is unlikely a confounder as strong as all the variables that were intendedly included in the model due to their confounding potential can exist. Yet, this criterion can be misleading whenever there is limited information on politicians' characteristics, insofar as many important characteristics could have been left out due to lack of information. Thus, evaluating sensibility based on a single strong benchmark and expert knowledge seems like a better strategy in these situations.

6 Closing remarks

Politician characteristic regression discontinuity designs are a popular strategy for evaluating the role of politician characteristics on an outcome of interest. Yet, as shown by Marshall (2022), the ability to retrieve causal effects from this methodology requires invoking unverifiable and likely unsuitable assumptions. Essentially, the problem rests in the fact that methodology does not compare winners to losers (say, outcomes for the politician who wins the election relative to outcomes of his runner-up), but winners of different types (a district managed by a politician with the outcome of interest who narrowly wins relative to one where a candidate without the trait narrowly wins). As a result, differences in other individual traits closely correlated to the trait of interest may confound the effect. More importantly, some of these differences will naturally emerge to keep elections close between opposing candidates. Thus, whenever possible, a researcher wishing to apply this methodology should attempt to describe the extent to which these potential sources of bias can alter the results.

This paper develops an econometric framework for addressing estimation bias in PCRD designs. First, I provide a covariate-adjusted local polynomial estimator that removes bias when all relevant confounders are observed. I then leverage the structure of this estimator to describe its asymptotic properties and to develop several decompositions of the bias. Furthermore, I argue that these decompositions are economically meaningful and can be used to understand other potential sources of bias.

In addition, I develop a procedure to evaluate the sensitivity of the results to unobserved confounders. To do so, I build on the work of Cinelli and Hazlett (2020a) to create bounds on the estimator that would result from controlling for the omitted characteristic. This procedure allows a researcher to assess the extent to which any unobserved characteristic could change their results. Moreover, by using other observed traits as benchmarks, the method allows the investigator to judge whether a confounder that can drastically change conclusions exists.

I apply all of these tools to a real-world example examining the impact of female leadership during the COVID-19 pandemic in Brazil (Bruce et al., 2022). This particular PCRD design exhibits multiple potential confounding factors, making it imperative to assess the potential size of the PCRD bias to draw any causal claims from the study. While observed unbalanced characteristics lead to biases, I show these are negligible on the aggregate due to emerging biases canceling each other out. Thus, I can conclude that observed differentials do not alter the qualitative effect found by the authors, reinforcing the claim of a positive effect due to female leadership. I also extend this analysis to account for the possibility of relevant omitted differentials explaining the results. Using the developed methods, I can show that unobserved confounding would need to be substantially large to overturn the main results, providing further reassurance on the existence of an effect driven solely by gender. This application exemplifies how the tools I developed can be useful in building credibility for any study leveraging a PCRD design to retrieve causal implications.

However, it is worth stressing that PCRD designs, like observational studies, do not provide the ideal environment for estimating causal effects. While the proposed methods provide a way to assess the confidence in making causality claims based on these studies, they do not surpass or substitute an appropriate identification strategy. The extent to which these methods are informative should always be gauged in light of the data quality used for covariate adjustment and bias benchmarking.

References

Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, *110*(3), 512–529. Retrieved from https://doi.org/10.1017/S0003055416000216

Alesina, A., Cassidy, T., & Troiano, U. (2019). Old and Young Politicians. *Economica*, *86*(344), 689–727. Retrieved 2023-05-01, from https://onlinelibrary.wiley.com/doi/abs/10.1111/ecca.12287 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecca.12287) doi: 10.1111/ecca.12287

Altonji, J., Elder, T., & Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151–184. Retrieved 2023-03-02, from https://www.jstor.org/stable/10.1086/426036 (Publisher: The University of Chicago Press) doi: 10.1086/426036

Beach, B., & Jones, D. B. (2017, February). Gridlock: Ethnic Diversity in Government and the Provision of Public Goods. *American Economic Journal: Economic Policy*, *9*(1), 112–136. Retrieved 2022-09-28, from https://www.aeaweb.org/articles?id=10.1257/pol.20150394 doi: 10.1257/pol.20150394

Besley, T., Montalvo, J. G., & Reynal-Querol, M. (2011). Do Educated Leaders Matter? *The Economic Journal*, *121*(554), F205–F227. Retrieved 2022-09-12, from http://www.jstor.org/stable/41237013 (Publisher: Wiley)

Bhalotra, S., & Clots-Figueras, I. (2014, May). Health and the Political Agency of Women. *American Economic Journal: Economic Policy*, *6*(2), 164–197. Retrieved 2023-02-18, from https://www.aeaweb.org/articles?id=10.1257/pol.6.2.164 doi: 10.1257/pol.6.2.164

Bhalotra, S., Clots-Figueras, I., Cassan, G., & Iyer, L. (2014, August). Religion, politician identity and development outcomes: Evidence from India. *Journal of Economic Behavior & Organization*, *104*, 4–17. Retrieved 2022-09-28, from https://www.sciencedirect.com/science/article/pii/S0167268113002345 doi: 10.1016/j.jebo.2013.09.006

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, *8*, 436–455.

Borusyak, K., & Hull, P. (2020, September). *Non-Random Exposure to Exogenous Shocks: Theory and Applications* [Working Paper]. National Bureau of Economic Research. Retrieved 2023-03-12, from https://www.nber.org/papers/w27845 doi: 10.3386/w27845

Bracco, E., Lockwood, B., Porcelli, F., & Redoano, M. (2015, March). Intergovernmental grants as signals and the alignment effect: Theory and evidence. *Journal of Public Economics*, *123*, 78–91. Retrieved 2023-02-18, from https://www.sciencedirect.com/science/article/pii/S0047272714002291 doi: 10.1016/j.jpubeco.2014.11.007

Bruce, R., Cavgias, A., Meloni, L., & Remígio, M. (2022, January). Under pressure: Women's leadership during the COVID-19 crisis. *Journal of Development Economics*, *154*, 102761. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/S0304387821001243 doi: 10.1016/j.jdeveco.2021.102761

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2018, April). On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association*, *113*(522), 767–779. Retrieved 2023-02-11, from https://doi.org/10.1080/01621459.2017.1285776 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2017.1285776) doi: 10.1080/01621459.2017.1285776

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019a, July). Regression Discontinuity Designs Using Covariates. *The Review of Economics and Statistics*, *101*(3), 442–451. Retrieved 2022-09-19, from https://doi.org/10.1162/rest_a_00760 doi: 10.1162/rest_a_00760

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019b, July). Regression Discontinuity Designs Using Covariates: Supplemental Appendix. *The Review of Economics and Statistics*.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals For Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. Retrieved 2022-09-12, from http://www.jstor.org/stable/43616914 (Publisher: [Wiley, The Econometric Society])

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015, October). Optimal Data-Driven Regression Discontinuity Plots. *Journal of the American Statistical Association*, *110*(512), 1753–1769. Retrieved 2023-02-20, from https://www.tandfonline.com/doi/full/10.1080/01621459.2015.1017578 doi: 10.1080/01621459.2015.1017578

Canon, D. T. (1993). Sacrificial Lambs or Strategic Politicians? Political Amateurs in U.S. House Elections. *American Journal of Political Science*, *37*(4), 1119–1141. Retrieved 2022-09-11, from https://www.jstor.org/stable/2111546 (Publisher: [Midwest Political Science Association, Wiley]) doi: 10.2307/2111546

Carson, J. L., Engstrom, E. J., & Roberts, J. M. (2007). Candidate Quality, the Personal Vote, and the Incumbency Advantage in Congress. *The American Political Science Review*, *101*(2), 289–301. Retrieved 2022-09-11, from https://www.jstor.org/stable/27644446 (Publisher: [American Political Science Association, Cambridge University Press])

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019, November). A Practical Introduction to Regression Discontinuity Designs: Foundations. *Elements in Quantitative and Computational Methods for the Social Sciences*. Retrieved 2023-02-17, from https://www.cambridge.org/core/elements/practical-introduction-to-regression-discontinuity-designs/F04907129D5C1B823E3DB19C31CAB905 (ISBN: 9781108684606 9781108710206 Publisher: Cambridge University Press) doi: 10.1017/9781108684606

Chemin, M. (2012). Welfare Effects of Criminal Politicians: A Discontinuity-Based Approach. *The Journal of Law & Economics*, *55*(3), 667–690. Retrieved 2023-02-18, from https://www.jstor.org/stable/10.1086/664574 (Publisher: [The University of Chicago Press, The Booth School of Business, University of Chicago, The University of Chicago Law School]) doi: 10.1086/664574

Cheng, M.-Y., Fan, J., & Marron, J. S. (1997, August). On automatic boundary corrections. *The Annals of Statistics*, *25*(4), 1691–1708. Retrieved 2023-02-08, from https://projecteuclid.org/journals/annals-of-statistics/volume-25/issue-4/On-automatic-boundary-corrections/10.1214/aos/1031594737.full (Publisher: Institute of Mathematical Statistics) doi: 10.1214/aos/1031594737

Cinelli, C., & Hazlett, C. (2020a). Making sense of sensitivity: extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *82*(1), 39–67. Retrieved 2023-02-18, from https://onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12348 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/rssb.12348) doi: 10.1111/rssb.12348

Cinelli, C., & Hazlett, C. (2020b). Online supplementary material for "Making Sense of Sensitivity: Extending Omitted Variable Bias". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *82*(1), 39–67. Retrieved 2023-02-18, from https://onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12348 doi: 10.1111/rssb.12348

Clots-Figueras, I. (2011, August). Women in politics: Evidence from the Indian States. *Journal of Public Economics*, *95*(7), 664–690. Retrieved 2022-09-12, from https://www.sciencedirect.com/science/article/pii/S0047272710001775 doi: 10.1016/j.jpubeco.2010.11.017

Dal Bó, E., Finan, F., Folke, O., Persson, T., & Rickne, J. (2017, November). Who Becomes A Politician? *The Quarterly Journal of Economics*, *132*(4), 1877–1914. Retrieved 2023-03-14, from https://doi.org/10.1093/qje/qjx016 doi: 10.1093/qje/qjx016

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., & Snyder, J. M. (2015). On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races. *American Journal of Political Science*, *59*(1), 259–274. Retrieved 2022-09-19, from https://www.jstor.org/stable/24363608 (Publisher: [Midwest Political Science Association, Wiley])

Elwert, F., & Winship, C. (2014, July). Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual review of sociology*, *40*, 31–53. Retrieved 2023-07-16, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6089543/ doi: 10.1146/annurev-soc-071913-043455

Ferreira, F., & Gyourko, J. (2014, April). Does gender matter for political leadership? The case of U.S. mayors. *Journal of Public Economics*, *112*, 24–39. Retrieved 2023-02-18, from https://www.sciencedirect.com/science/article/pii/S0047272714000073 doi: 10.1016/j.jpubeco.2014.01.006

Fortin, N., Lemieux, T., & Firpo, S. (2011, January). Chapter 1 - Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4, pp. 1–102). Elsevier. Retrieved 2023-03-26, from https://www.sciencedirect.com/science/article/pii/S0169721811004072 doi: 10.1016/S0169-7218(11)00407-2

Fowler, A., & Hall, A. B. (2015, December). Congressional seniority and pork: A pig fat myth? *European Journal of Political Economy*, *40*, 42–56. Retrieved 2023-02-18, from https://linkinghub.elsevier.com/retrieve/pii/S0176268015000774 doi: 10.1016/j.ejpoleco.2015.07.006

Frölich, M., & Huber, M. (2019a). Including Covariates in the Regression Discontinuity Design. *Journal of Business & Economic Statistics*, *37*(4), 736–748. Retrieved from https://doi.org/10.1080/07350015.2017.1421544 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2017.1421544) doi: 10.1080/07350015.2017.1421544

Frölich, M., & Huber, M. (2019b). Including Covariates in the Regression Discontinuity Design: Supplemental Material. *Journal of Business & Economic Statistics*, *37*(4), 736–748. Retrieved from https://doi.org/10.1080/07350015.2017.1421544 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2017.1421544) doi: 10.1080/07350015.2017.1421544

Gagliarducci, S., & Paserman, M. D. (2012). Gender Interactions within Hierarchies: Evidence from the Political Arena. *The Review of Economic Studies*, *79*(3), 1021–1052. Retrieved 2023-03-31, from https://www.jstor.org/stable/23261378 (Publisher: [Oxford University Press, Review of Economic Studies, Ltd.])

Gelman, A., & Imbens, G. (2019, July). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456. Retrieved 2022-09-19, from https://doi.org/10.1080/07350015.2017.1366909 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2017.1366909) doi: 10.1080/07350015.2017.1366909

Goldsmith-Pinkham, P., Hull, P., & Kolesár, M. (2022, June). *Contamination Bias in Linear Regressions* [Working Paper]. National Bureau of Economic Research. Retrieved 2023-04-01, from https://www.nber.org/papers/w30108 doi: 10.3386/w30108

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209. Retrieved 2023-02-08, from https://www.jstor.org/stable/2692190 (Publisher: [Wiley, Econometric Society])

Hall, A. B. (2015, February). What Happens When Extremists Win Primaries? *American Political Science Review*, *109*(1), 18–42. Retrieved 2023-02-18, from https://www.cambridge.org/core/journals/american-political-science-review/article/what-happens-when-extremists-win-primaries/ECAC69648AE0DF91D93103E18342B9D2 (Publisher: Cambridge University Press) doi: 10.1017/S0003055414000641

Hopkins, D. J., & McCabe, K. T. (2012, July). After It's Too Late: Estimating the Policy Impacts of Black Mayoralties in U.S. Cities. *American Politics Research*, *40*(4), 665–700. Retrieved 2022-09-28, from https://doi.org/10.1177/1532673X11432469 (Publisher: SAGE Publications Inc) doi: 10.1177/1532673X11432469

Hosman, C. A., Hansen, B. B., & Holland, P. W. (2010). The Sensitivity of Linear Regression Coefficients' Confidence Limits to the Omission of a Confounder. *The Annals of Applied Statistics*, *4*(2), 849–870. Retrieved 2023-03-02, from https://www.jstor.org/stable/29765533 (Publisher: Institute of Mathematical Statistics)

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., & Tukiainen, J. (2018). When does regression discontinuity design work? Evidence from random election outcomes. *Quantitative Economics*, *9*(2), 1019–1051. Retrieved 2022-09-19, from https://onlinelibrary.wiley.com/doi/abs/10.3982/QE864 doi: 10.3982/QE864

Imbens, G. (2003, May). Sensitivity to Exogeneity Assumptions in Program Evaluation. *American Economic Review*, *93*(2), 126–132. Retrieved 2023-03-02, from https://www.aeaweb.org/articles?id=10.1257/000282803321946921 doi: 10.1257/000282803321946921

Imbens, G., & Kalyanaraman, K. (2012, July). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, *79*(3), 933–959. Retrieved 2023-02-11, from https://

doi.org/10.1093/restud/rdr043 doi: 10.1093/restud/rdr043

Imbens, G., & Wooldridge, J. M. (2009, March). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86. Retrieved 2023-04-01, from https://www.aeaweb.org/articles?id=10.1257/jel.47.1.5 doi: 10.1257/jel.47.1.5

Jones, M. P., Saiegh, S., Spiller, P. T., & Tommasi, M. (2002). Amateur Legislators – Professional Politicians: The Consequences of Party-Centered Electoral Rules in a Federal System. *American Journal of Political Science*, *46*(3), 656–669. Retrieved 2022-09-11, from https://www.jstor.org/stable/3088406 (Publisher: [Midwest Political Science Association, Wiley]) doi: 10.2307/3088406

Kirkland, P. A. (2021, October). Business Owners and Executives as Politicians: The Effect on Public Policy. *The Journal of Politics*, *83*(4), 1652–1668. Retrieved 2023-02-13, from https://www.journals.uchicago.edu/doi/abs/10.1086/715067 (Publisher: The University of Chicago Press) doi: 10.1086/715067

Kitagawa, E. M. (1955, December). Components of a Difference Between Two Rates. *Journal of the American Statistical Association*, *50*(272), 1168–1194. Retrieved 2023-02-28, from https://doi.org/10.1080/01621459.1955.10501299 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.1955.10501299) doi: 10.1080/01621459.1955.10501299

Lawless, J. L. (2015). Female Candidates and Legislators. *Annual Review of Political Science*, *18*(1), 349–366. Retrieved 2023-04-03, from https://doi.org/10.1146/annurev-polisci-020614-094613 (_eprint: https://doi.org/10.1146/annurev-polisci-020614-094613) doi: 10.1146/annurev-polisci-020614-094613

Lee, D. S. (2008, February). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697. Retrieved 2022-09-12, from https://www.sciencedirect.com/science/article/pii/S0304407607001121 doi: 10.1016/j.jeconom.2007.05.004

Lewis, B. D., Nguyen, H. T. M., & Hendrawan, A. (2020, September). Political accountability and public service delivery in decentralized Indonesia: Incumbency advantage and the performance of second term mayors. *European Journal of Political Economy*, *64*, 101910. Retrieved 2023-02-18, from https://www.sciencedirect.com/science/article/pii/S0176268020300586 doi: 10.1016/j.ejpoleco.2020.101910

Marshall, J. (2022). Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics? *American Journal of Political Science*, *n/a*(n/a). Retrieved 2022-09-19, from https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12741 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12741) doi: 10.1111/ajps.12741

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It. *American Journal of Political Science*, *62*(3), 760–775. Retrieved 2023-07-12, from https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12357 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12357) doi: 10.1111/ajps.12357

Mukherjee, D., Banerjee, M., & Ritov, Y. (2021, November). *Estimation of a score-explained non-*

randomized treatment effect in fixed and high dimensions. arXiv. Retrieved 2023-02-24, from http://arxiv.org/abs/2102.11229 (arXiv:2102.11229 [math, stat]) doi: 10.48550/arXiv.2102.11229

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, *14*(3), 693–709. Retrieved 2023-02-28, from https://www.jstor.org/stable/2525981 (Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University]) doi: 10.2307/2525981

Oster, E. (2019, April). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. Retrieved 2023-03-02, from https://doi.org/10.1080/07350015.2016.1227711 (Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2016.1227711) doi: 10.1080/07350015.2016.1227711

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge: Cambridge University Press. Retrieved 2023-07-16, from https://www.cambridge.org/core/books/causality/B0046844FAE10CBF274D4ACBDAEB5F5B doi: 10.1017/CBO9780511803161

Peng, S., & Ning, Y. (2019, November). *Regression Discontinuity Design under Self-selection* (Papers No. 1911.09248). arXiv.org. Retrieved from https://ideas.repec.org/p/arx/papers/1911.09248.html

Pettersson-Lidbom, P. (2008). Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach. *Journal of the European Economic Association*, *6*(5), 1037–1056. Retrieved 2023-02-18, from https://onlinelibrary.wiley.com/doi/abs/10.1162/JEEA.2008.6.5.1037 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1162/JEEA.2008.6.5.1037) doi: 10.1162/JEEA.2008.6.5.1037

Power, T. J., & Rodrigues-Silveira, R. (2019, February). Mapping Ideological Preferences in Brazilian Elections, 1994-2018: A Municipal-Level Study. *Brazilian Political Science Review*, *13*, e0001. Retrieved 2023-04-23, from http://www.scielo.br/j/bpsr/a/g9ZMS7VwNWSm6QPQqjNSMgk/?lang=en (Publisher: Associação Brasileira de Ciência Política) doi: 10.1590/1981-3821201900010001

Robins, M. A. H., James M. (2023). *Causal Inference: What If.* Boca Raton: CRC Press.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer Science & Business Media.

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome. *Journal of the Royal Statistical Society. Series B (Methodological)*, *45*(2), 212–218. Retrieved 2023-03-02, from https://www.jstor.org/stable/2345524 (Publisher: [Royal Statistical Society, Wiley])

Szakonyi, D. (2021, January). Private Sector Policy Making: Business Background and Politicians' Behavior in Office. *The Journal of Politics*, *83*(1), 260–276. Retrieved 2023-02-13, from https://www.journals.uchicago.edu/doi/abs/10.1086/709297 (Publisher: The University of Chicago Press) doi: 10.1086/709297

van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge, UK ; New York, NY, USA: Cambridge University Press.

Wooldridge, J. M. (2021, August). *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and*

Difference-in-Differences Estimators [SSRN Scholarly Paper]. Rochester, NY. Retrieved 2023-04-01, from https://papers.ssrn.com/abstract=3906345 doi: 10.2139/ssrn.3906345

Xu, Y., & Yao, Y. (2015, May). Informal Institutions, Collective Action, and Public Investment in Rural China. *American Political Science Review*, *109*(2), 371–391. Retrieved from https://www.cambridge.org/core/journals/american-political-science-review/article/abs/informal-institutions-collective-action-and-public-investment-in-rural-china/D0C7916127704E4D9C652D36A8525D8D (Publisher: Cambridge University Press) doi: 10.1017/S0003055415000155

Appendix for:

"Addressing Bias in Politician Characteristic Regression Discontinuity Designs"

Contents

A Appendix - Monte Carlo Simulation

In this section, I evaluate the performance of the proposed estimator $\hat{\tau}^{BC}_{PCRD}$ with respect to the traditional $\hat{\tau}_{PCRD}$ estimator. To do this, I recreate a similar situation to Marshall (2022, p.9), describing a typical data generation process leading to a PCRD bias.

Consider candidates i and j running in the same district d. Each of the politicians is either type $X_{id} = 1$ or $X_{id} = 0$, with $X_{id} \sim \text{Bernoulli}(0.5)$. Moreover, these candidates have two potential compensating differentials $\mathbf{Z}_{id} = (Z_{id1}, Z_{id2})$. The distribution of these traits depends on the type of politician and one another, thus exhibiting mutual correlation and leading to a PCRD estimation bias. In particular:

$$Z_{id1} \sim \begin{cases} \exp(1) & \text{if } X_{id} = 0, \\ \exp(2) & \text{if } X_{id} = 1 \end{cases}$$

and

$$Z_{id2} \sim \begin{cases} \text{Binomial}(5, 0.5) & \text{if } Z_{id1} > 1, \\ \text{Binomial}(3, 0.4) & \text{if } Z_{id1} \leq 1 \end{cases}$$

Moreover, the votes obtained by each candidate are a function of his attributes. More precisely,

$$V_{id} = \alpha \left(\frac{X_{id} - X_{jd}}{2} \right) + \beta_1 \left(\frac{Z_{id1} - Z_{jd1}}{2} \right) + \beta_2 \left(\frac{Z_{id2} - Z_{jd2}}{2} \right) + \left(\frac{\xi_{id} - \xi_{jd}}{2} \right)$$

where $\alpha = 0.3$, $\beta_1 = 5$, $\beta_2 = 7$ and $\xi_{id} \sim \mathcal{N}(0, 0.1)$. Hence,

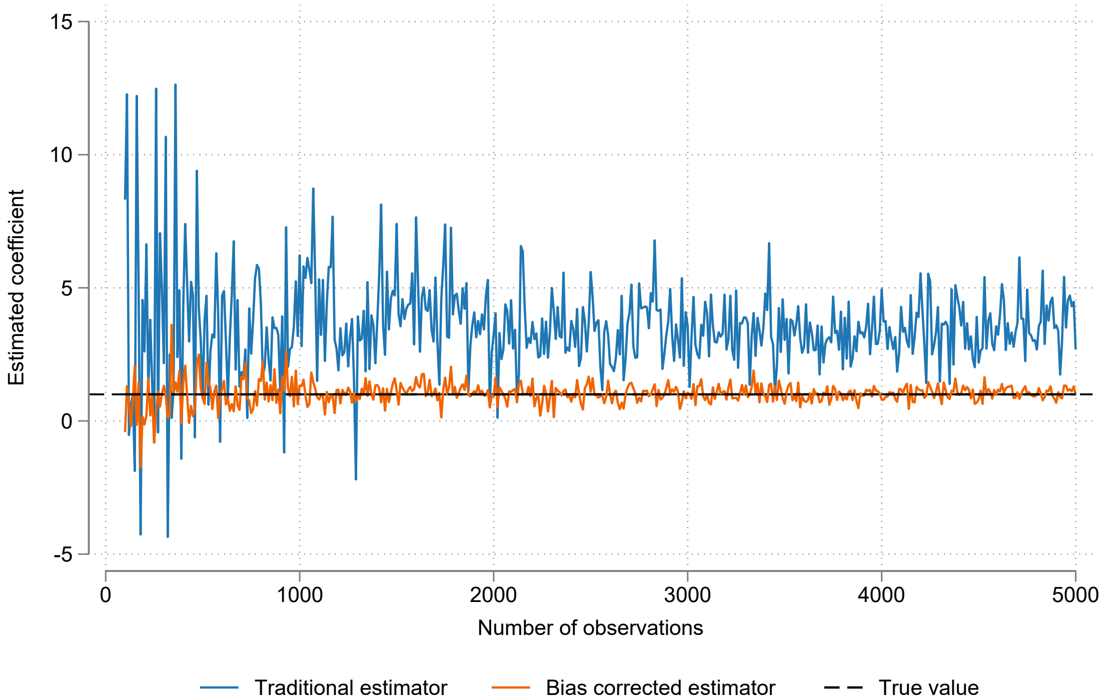
$$\Delta_d = \alpha + \beta_1 (Z_{1d1} - Z_{0d1}) + \beta_2 (Z_{1d2} - Z_{0d2}) + (\xi_{1d} - \xi_{0d})$$

Moreover, the data generation process is non-linear and includes interaction between these characteristics:

$$Y_{id}(1) = 1 \cdot X_{id} + 5 \log(1 + Z_{id1}) * Z_{id2} + Z^2_{id2} + \varepsilon_d$$

where $\varepsilon_d \sim \mathcal{N}(0, 0.3)$ is an i.i.d district-level shock. Following this model, the value to be estimated is $\tau_{PCRD} = 1$.

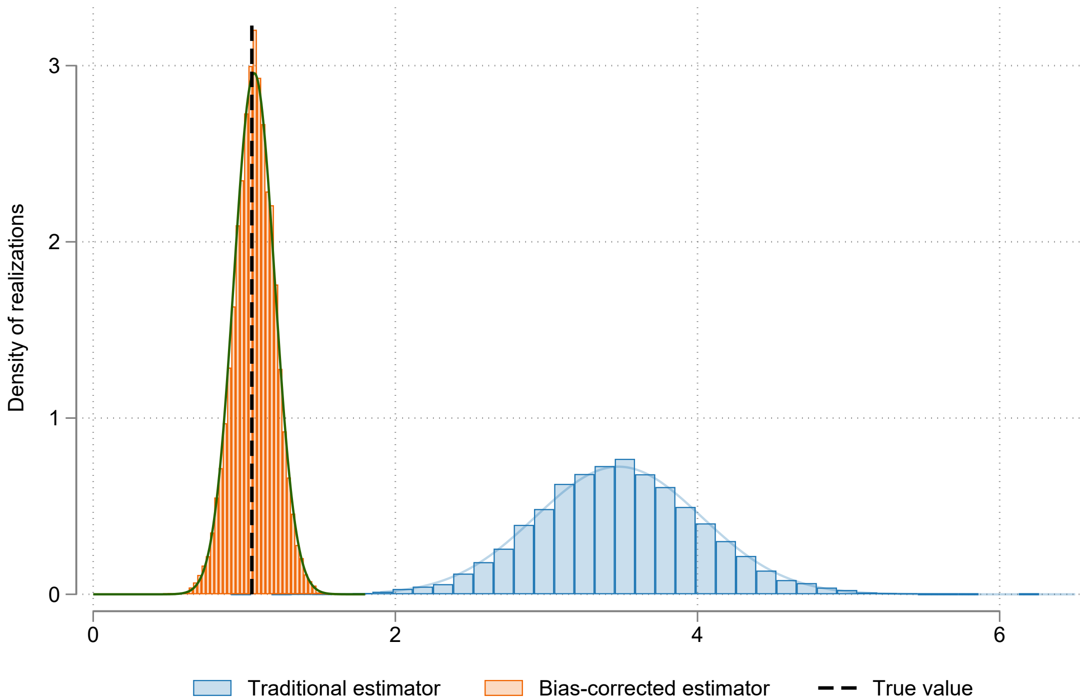
FIGURE A-1: *Monte Carlo simulation to test the consistency of the bias-corrected estimator.*



Notes: Traditional estimator refers to a local linear model, with triangular kernel and MSE-optimal bandwidths following Calonico et al. (2014).

– As it can be observed in Figure A-1, while $\hat{\tau}^{BC}_{PCRD}$ accurately estimates the correct value, $\hat{\tau}_{PCRD}$ is consistently overestimating the effect. This is because the existence of unbalanced characteristics in close elections causes the PCRD estimation bias. Furthermore, not only does the PCRD consistently estimate the effect, but its variance is significantly lower than that of the traditional estimator. Hence, the proposed estimator greatly improves the precision with which τ_{PCRD} is estimated. A similar exercise portrayed in Figure A-2, is consistent with the estimator's theoretical asymptotic normality and variance reduction properties (since in this case $\hat{\hat{\gamma}}_+ - \hat{\hat{\gamma}}_- \xrightarrow{\mathbb{P}} 0$).

FIGURE A-2: *Monte Carlo simulation to test asymptotic normality.*



Notes: Fix sample size of $n = 10000$. The simulation portrays the results of 10000 random realizations for each estimator. Traditional estimator refers to a local linear model, with triangular kernel and MSE-optimal bandwidths following Calonico et al. (2014).

B Appendix - Technical Information and Ommited Proofs

B.1 Basic notation

A rigorous treatment of the results presented in the paper demands dealing with the theory of local polynomial estimators. In this section, I will briefly list the basic expressions that will be required for accurately enunciating assumptions and proving relevant results. In most cases, I borrow the notation from Calonico et al. (2019a).

Given the notation introduced in the text, define the following vectors and matrices:

$$
\begin{aligned}
&\mathbf{Y} = [Y_1, \cdots, Y_n]', && \Delta = [\Delta_1, \cdots, \Delta_n]', \\
&\mathbf{Z} = [\mathbf{Z}_1, \cdots, \mathbf{Z}_n]', && \mathbf{Z}_d = [Z_{d1}, Z_{d2}, \cdots, Z_{dm}]', \quad d = 1, 2, \cdots, n, \\
&\mathbf{Y}(0) = [Y_1(0), \cdots, Y_n(0)]', && \mathbf{Y}(1) = [Y_1(1), \cdots, Y_n(1)]', \\
&\mathbf{Z}(0) = [\mathbf{Z}_1(0), \cdots, \mathbf{Z}_n(0)]', && \mathbf{Z}(1) = [\mathbf{Z}_1(1), \cdots, \mathbf{Z}_n(1)]', \\
&\mu_{Y-}(\Delta) = \mathbb{E}[\mathbf{Y}(0) \mid \Delta], && \mu_{Y+}(\Delta) = \mathbb{E}[\mathbf{Y}(1) \mid \Delta], \\
&\Sigma_{Y-} = \mathbb{V}[\mathbf{Y}(0) \mid \Delta], && \Sigma_{Y+} = \mathbb{V}[\mathbf{Y}(1) \mid \Delta], \\
&\mu_{Z-}(\Delta) = \mathbb{E}[\mathrm{vec}(\mathbf{Z}(0)) \mid \Delta], && \mu_{Z+}(\Delta) = \mathbb{E}[\mathrm{vec}(\mathbf{Z}(1)) \mid \Delta]. \\
&\Sigma_{Z-} = \mathbb{V}[\mathrm{vec}(\mathbf{Z}(0)) \mid \Delta], && \Sigma_{Z+} = \mathbb{V}[\mathrm{vec}(\mathbf{Z}(1)) \mid \Delta].
\end{aligned}
$$

Recall from the text that:

$$
\begin{aligned}
\mu_{Y-}(u) &= \mathbb{E}\left[Y_d(0) \mid \Delta_d = u\right], & \mu_{Y+}(u) &= \mathbb{E}\left[Y_d(1) \mid \Delta_d = u\right], \\
\sigma^2_{Y-}(u) &= \mathbb{V}\left[Y_d(0) \mid \Delta_d = u\right], & \sigma^2_{Y+}(u) &= \mathbb{V}\left[Y_d(1) \mid \Delta_d = u\right], \\
\mu_{Z-}(u) &= \mathbb{E}\left[\mathbf{Z}_d(0) \mid \Delta_d = u\right], & \mu_{Z+}(u) &= \mathbb{E}\left[\mathbf{Z}_d(1) \mid \Delta_d = u\right], \\
\sigma^2_{Z-}(u) &= \mathbb{V}\left[\mathbf{Z}_d(0) \mid \Delta_d = u\right], & \sigma^2_{Z+}(u) &= \mathbb{V}\left[\mathbf{Z}_d(1) \mid \Delta_d = u\right],
\end{aligned}
$$

where

$$
\mu_{Z_{\ell-}}(u) = \mathbb{E}\left[Z_{d\ell}(0) \mid \Delta_d = u\right], \quad \mu_{Z_{\ell}+}(u) = \mathbb{E}\left[Z_{d\ell}(1) \mid \Delta_d = u\right],
$$

for $\ell = 1, 2, \cdots, m$.

Additionally, handling the joint distribution of the outcome variable and the additional covariates becomes necessary whenever estimation involves covariates. Thus, consider the following:

$$
\begin{aligned}
&\mathbf{S}_d = \left[Y_d, \mathbf{Z}_d'\right]', && \mathbf{S}_d(0) = [Y_d(0), \mathbf{Z}_d(0)']', \quad \mathbf{S}_d(1) = [Y_d(1), \mathbf{Z}_d(1)']', \\
&\mathbf{S} = [\mathbf{Y}, \mathbf{Z}], && \mathbf{S}(0) = [\mathbf{Y}(0), \mathbf{Z}(0)], \quad \mathbf{S}(1) = [\mathbf{Y}(1), \mathbf{Z}(1)], \\
&\mu_{S_-}(\Delta) = \mathbb{E}[\mathrm{vec}(\mathbf{S}(0)) \mid \Delta], && \mu_{S+}(\Delta) = \mathbb{E}[\mathrm{vec}(\mathbf{S}(1)) \mid \Delta], \\
&\Sigma_{S-} = \mathbb{V}[\mathrm{vec}(\mathbf{S}(0)) \mid \Delta], && \Sigma_{S_+} = \mathbb{V}[\mathrm{vec}(\mathbf{S}(1)) \mid \Delta], \\
&\mu_{S_-}(u) = \mathbb{E}\left[\mathbf{S}_d(0) \mid \Delta_d = u\right], && \mu_{S+}(u) = \mathbb{E}\left[\mathbf{S}_d(1) \mid \Delta_d = u\right], \\
&\sigma^2_{S_-}(u) = \mathbb{V}\left[\mathbf{S}_d(0) \mid \Delta_d = u\right], && \sigma^2_{S+}(u) = \mathbb{V}\left[\mathbf{S}_d(1) \mid \Delta_d = u\right].
\end{aligned}
$$

Moreover, consider the shorthand notation

$$\Gamma_{-,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_-(h)\mathbf{R}_p(h)/n, \quad \Upsilon_{Y-,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_-(h)\mathbf{Y}/n$$
$$\Gamma_{+,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_+(h)\mathbf{R}_p(h)/n, \quad \Upsilon_{Y+,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_+(h)\mathbf{Y}/n$$

where

$$\mathbf{R}_p(h) = \left[\mathbf{r}_p\left(\frac{\Delta_1}{h}\right), \mathbf{r}_p\left(\frac{\Delta_2}{h}\right), \cdots, \mathbf{r}_p\left(\frac{\Delta_n}{h}\right)\right]'_{n\times(1+p)},$$
$$\mathbf{H}_p(h) = \text{diag}\left(h^j : j = 0,1,\cdots,p\right) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & h & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h^p \end{bmatrix}_{(1+p)\times(1+p)}$$
$$\mathbf{K}_-(h) = \text{diag}\left(\mathbb{1}\left(\Delta_d < 0\right) k_h\left(-\Delta_d\right) : d = 1,2,\cdots,n\right),$$
$$\mathbf{K}_+(h) = \text{diag}\left(\mathbb{1}\left(\Delta_d \geq 0\right) k_h\left(\Delta_d\right) : d = 1,2,\cdots,n\right).$$

and $\mathbf{r}_p(x) = (1, x, \cdots, x^p)'$, $k_h(u) = K(u/h)/h$, with $K(\cdot)$ a fixed kernel function, together with h a positive bandwidth sequence.

Finally, to simplify notation, set

$$\mathbf{P}_{-,p}(h) = \sqrt{h}\Gamma^{-1}_{-,p}(h)\mathbf{R}_p(h)'\mathbf{K}_-(h)/\sqrt{n},$$
$$\mathbf{P}_{+,p}(h) = \sqrt{h}\Gamma^{-1}_{+,p}(h)\mathbf{R}_p(h)'\mathbf{K}_+(h)/\sqrt{n},$$

and, with a slight abuse of notation $\mathbf{v}^k = (v_1^k, v_2^k, \cdots, v_n^k)'$ for $\mathbf{v} \in \mathbb{R}^n$,

$$\vartheta_{-,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_-(h)(\Delta/n)^{1+p}/n,$$
$$\vartheta_{+,p}(h) = \mathbf{R}_p(h)'\mathbf{K}_-(h)(\Delta/n)^{1+p}/n$$

B.2 Regularity assumptions

Local polynomial estimators require continuity or smoothness assumptions to ensure desirable properties such as consistency and asymptotic normality. For this paper, I use the standard conditions presented by Calonico et al. (2019b).

Assumption RA-1 (Kernel)**.** *For some $\zeta > 0$, the kernel function $K : [0, \zeta] \to \mathbb{R}$ is bounded and non-negative, zero outside its support, and positive and continuous on $(0, \zeta)$.*

Assumption RA-1 limits the number of valid kernels that can be used to localize the effect around the cutoff. Popular kernels, such as the triangular, uniform, and Epanechnikov's, are allowed.

Assumption RA-2 (RD standard)**.** *For some $\zeta_0 > 0$, the following hold in some neighborhood of 0 for all $u \in (-\zeta_0, \zeta_0)$:*

a) The Lebesgue density of Δ_d, denoted by $f_\Delta(u)$, is continuous and bounded away from zero.

b) $\mathbb{E}[|Y_d(t)|^4|\Delta_d = u]$ *are continuous for* $t \in \{0,1\}$

c) $\mu_{Y-}(u)$ *and* $\mu_{Y+}(u)$ *are* $\rho \geq 1$ *times continuously differentiable.*

d) $\sigma_{Y-}(u)$ *and* $\sigma_{Y+}(u)$ *are continuous and bounded away from zero.*

Assumption RA-2 describes the minimum regularity conditions needed to estimate the effects correctly. Part (a) requires the running variable to be continuously distributed near the cutoff and guarantees the presence of observations arbitrarily close to the threshold in large samples. Part (b) imposes the existence of fourth moments. Parts (c) and (d) are standard smoothness conditions on the underlying regression functions and conditional variances. Assumption RA-3 requires analogous properties for the joint distribution of the outcome variable and the covariates.

Assumption RA-3 (RD covariates)**.** *For some* $\zeta_0 > 0$*, the following hold in some neighborhood of* 0 *for all* $u \in (-\zeta_0, \zeta_0)$:

a) $\mathbb{E}\left[\mathbf{Z}_d(0)Y_d(0) \mid \Delta_d = u\right]$ *and* $\mathbb{E}\left[\mathbf{Z}_d(1)Y_d(1) \mid \Delta_d = u\right]$ *are continuously differentiable.*

b) $\mu_{S-}(u)$ *and* $\mu_{S+}(u)$ *are* $\rho \geq 1$ *times continuously differentiable.*

c) $\sigma^2_{S-}(u)$ *and* $\sigma^2_{S+}(u)$ *are continuous and invertible.*

d) $\mathbb{E}\left[|\mathbf{S}_d(t)|^4 \mid X_i = x\right], t \in \{0,1\}$*, are continuous.*

where $|\cdot|$ *denotes the Euclidean matrix norm, that is,* $|A|^2 =$ *trace*$(A'A)$ *for scalar, vector, or matrix A.*

Additionally, non-parametric identification requires additional regularity assumptions. I collect some of the conditions specified in Frölich and Huber (2019a) required for identification and estimation in the sharp RD design.

Assumption RA-4 (Non-parametric identification regularity conditions)**.** *For a given* $\eta > 0$*, let* $U^+_\eta = [0, \eta]$ *and* $U^-_\eta = [-\eta, 0]$ *and* $U_\eta = U^+_\eta \cup U^-_\eta$*. Then*

a) The cumulative distribution function of the running variable $F_\Delta(u)$ *is differentiable at* $u = 0$ *and* $f_\Delta(0) > 0$.

b) The conditional cumulative distribution functions

$$\lim_{\eta \to 0} F_{\mathbf{Z}_d|\Delta_d \in U^+_\eta}(z) \qquad \text{and} \qquad \lim_{\eta \to 0} F_{\mathbf{Z}_d|\Delta_d \in U^-_\eta}(z)$$

exist.

c) There exist probability density/mass functions $f^+_{\mathbf{Z}_d(1)|\Delta_d}(z|0)$ *and* $f^-_{\mathbf{Z}_d(0)|\Delta_d}(z|0)$ *linked to the random vectors* $\mathbf{Z}_d(1)|\Delta_d = 0$ *and* $\mathbf{Z}_d(0)|\Delta_d = 0$ *respectively.*

d) *There exists* $\eta_0 > 0$ *such that regression functions*

$$\mathbb{E}[Y_d(1)|\mathbf{Z}_d(1) = z, \Delta_d = u] \qquad \text{and} \qquad \mathbb{E}[Y_d(0)|\mathbf{Z}_d(1) = z, \Delta_d = u]$$

are bounded almost surely over U_η*, for any* $\eta < \eta_0$.

Finally, I introduce regularity assumptions for the estimation in the general case.

Assumption RA-5 (Non-parametric estimation RD conditions)**.** *For a given* $\eta > 0$*, let* $U_\eta^+ = [0, \eta]$ *and* $U_\eta^- = [-\eta, 0]$ *and* $U_\eta = U_\eta^+ \cup U_\eta^-$*. Then*

a) $\bar{K}(\cdot)$ *and* $\tilde{K}(\cdot)$ *are kernels of order* $\lambda \geq 2$*, satisfying all conditions in Assumption RA-1.*

b) $m^+(z,u)$ *and* $m^-(z,u)$ *are* λ *times continuously differentable with respect to u at* $u = 0$*, with the* $\lambda-$*th derivative Hölder continuous in a neighborhood of* $u = 0$.

c) *The probability density functions* $f^+_{(\mathbf{Z}(\mathbf{1}),\Delta_d)}(z,u)$ *and* $f^-_{(\mathbf{Z}(\mathbf{0}),\Delta_d)}(z,u)$ *are* $\lambda - 1$ *times continuously differentiable with respect to u at* $u = 0$*, with the* $(\lambda - 1)-$*th derivative Hölder continuous in a neighborhood of* $u = 0$.

d) $m^+(z,u)$ *and* $f^+_{(\mathbf{Z}(\mathbf{1}),\Delta_d)}(z,u)$ *have two continuous right derivatives with respect to z at* $u = 0$*, with second derivative Hölder continuous in a neighborhood of* $u = 0$.

e) $m^-(z,u)$ *and* $f^-_{(\mathbf{Z}(\mathbf{0}),\Delta_d)}(z,u)$ *have two continuous left derivatives with respect to z at* $u = 0$*, with second derivative Hölder continuous in a neighborhood of* $u = 0$.

f) *The bandwidths* $h, h_z, \bar{h} \to 0$*,* $nh \to \infty$*,* $nh \to \infty$ *and* $nh_z h_x^{m_c} \to \infty$*, with* m_c *the number of continuous covariates.*

g) *The left and right limits of the conditional variances*

$$\lim_{\eta \to 0} \mathbb{E}[(Y - m^+(\mathbf{Z}_d, \Delta_d))^2 | \mathbf{Z}_d, \Delta_d = u + \eta] \quad \text{and} \quad \lim_{\eta \to 0} \mathbb{E}[(Y - m^-(\mathbf{Z}_d, \Delta_d))^2 | \mathbf{Z}_d, \Delta_d = u - \eta]$$

exist at $u = 0$.

B.3 Standard local polynomial estimators

In this section, I will explain briefly how local polynomial estimators work. For a given degree p, the standard local model for the dependent variable Y_d can be formulated as:

$$\tilde{Y}_d = r_p(\Delta_d)\tilde{\beta}_{Y-,p} + (X_d \times r_p(\Delta_d))\tilde{\beta}_{Y+,p}$$

Localization requires estimating the model in a neighborhood of the threshold. Thus, consider bandwidths sequences $\mathbf{h} = (h_-, h_+)$ used for restricting the sample at the left and right of the cutoff respectively, and a weighting scheme $K_{\mathbf{h}}(u) = \mathbb{1}(u < 0)k_{h-}(-u) + \mathbb{1}(u \geq 0)k_{h+}(u)$.

The target population parameters are the coefficients of the Taylor expansion of the conditional expectation function around the threshold, namely:

$$\beta_{Y-,p} = \left[\mu_{Y-}(0),\ \frac{1}{1!}\mu_{Y-}^{(1)}(0),\ \frac{1}{2!}\mu_{Y-}^{(2)}(0)\cdots,\ \frac{1}{p!}\mu_{Y-}^{(p)}(0)\right]'_{(1+p)\times 1},$$
$$\beta_{Y+,p} = \left[\mu_{Y+}(0),\ \frac{1}{1!}\mu_{Y+}^{(1)}(0),\ \frac{1}{2!}\mu_{Y+}^{(2)}(0)\cdots,\ \frac{1}{p!}\mu_{Y+}^{(p)}(0)\right]'_{(1+p)\times 1}$$

where, for every $\nu = 0, 1, \cdots,$

$$\mu_{Y-}^{(\nu)}(u) = \frac{\partial^\nu}{\partial u^\nu}\mu_{Y-}(u) = \frac{\partial^\nu}{\partial u^\nu}\mathbb{E}[Y_d(0)|\Delta_d = u], \quad \mu_{Y+}^{(\nu)}(u) = \frac{\partial^\nu}{\partial u^\nu}\mu_{Y+}(u) = \frac{\partial^\nu}{\partial u^\nu}\mathbb{E}[Y_d(1)|\Delta_d = u].$$

The local polynomial estimators result from solving a weighted least square problem:

$$\hat{\hat{\beta}}_{Y,p}(\mathbf{h}) = \left[\begin{array}{c} \hat{\hat{\beta}}_{Y-,p}(h-) \\ \hat{\hat{\beta}}_{Y+,p}(h+) \end{array}\right]_{(2+2p)} = \underset{b_-,b_+}{argmin} \sum_{d=1}^{n} \left(Y_d - \mathbf{r}_{-,p}(\Delta_d)'\mathbf{b}_- - \mathbf{r}_{+,p}(\Delta_d)'\mathbf{b}_+\right)^2 K_{\mathbf{h}}(\Delta_d)$$

where $\mathbf{b}_- \in \mathbb{R}^{1+p}, \mathbf{b}_+ \in \mathbb{R}^{1+p}$, and

$$\mathbf{r}_{-,p}(u) := \mathbb{1}(u < 0)\mathbf{r}_p(u), \quad \mathbf{r}_{+,p}(u) := \mathbb{1}(u \geq 0)\mathbf{r}_p(u).$$

To express the problem in matrix form, consider the following auxiliary matrices:

$$\begin{aligned}
\tilde{\mathbf{R}}_{-,p}(h) &= \left[\mathbb{1}(\Delta_1 < 0)\mathbf{r}_p\left(\frac{\Delta_1}{h}\right), \cdots, \mathbb{1}(\Delta_n < 0)\mathbf{r}_p\left(\frac{\Delta_n}{h}\right)\right]'_{n\times(p+1)} \\
\tilde{\mathbf{R}}_{+,p}(h) &= \left[\mathbb{1}(\Delta_1 \geq 0)\mathbf{r}_p\left(\frac{\Delta_1}{h}\right), \cdots, \mathbb{1}(\Delta_n \geq 0)\mathbf{r}_p\left(\frac{\Delta_n}{h}\right)\right]'_{n\times(p+1)} \\
\tilde{\mathbf{R}}_p(\mathbf{h}) &= \left(\tilde{\mathbf{R}}_{-,p}(h_-)\mathbf{H}_p^{-1}(h_-) \quad \tilde{\mathbf{R}}_{+,p}(h_+)\mathbf{H}_p^{-1}(h_+)\right)_{n\times 2(p+1)} \\
\mathbf{W}(\mathbf{h}) &= (\mathbf{K}_-(h_-) + \mathbf{K}_+(h_+))_{n\times n}
\end{aligned}$$

It follows then that

$$\hat{\hat{\beta}}_{Y,p}(\mathbf{h}) = \underset{b}{argmin}\ (\mathbf{Y} - \tilde{\mathbf{R}}_p(\mathbf{h})\mathbf{b})'\mathbf{W}(\mathbf{h})(\mathbf{Y} - \tilde{\mathbf{R}}_p(\mathbf{h})\mathbf{b})$$

where $\mathbf{b} = (\mathbf{b}_- \quad \mathbf{b}_+)' \in \mathbb{R}^{2+2p}$.

Attempting to solve this problem leads to the first-order condition

$$\hat{\hat{\beta}}_{Y,p}(\mathbf{h}) = (\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}))^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y}$$

which results in

$$\begin{pmatrix}\hat{\hat{\beta}}_{Y-,p}(h-)\\ \hat{\hat{\beta}}_{Y+,p}(h+)\end{pmatrix} = \begin{pmatrix}([\mathbf{R}_p(h_-)\mathbf{H}_p^{-1}(h_-)]'\mathbf{K}_-(h_-)\mathbf{R}_p(h_-)\mathbf{H}_p^{-1}(h_-))^{-1}[\mathbf{R}_p(h_-)\mathbf{H}_p^{-1}(h_-)]'\mathbf{K}_-(h_-)\mathbf{Y}\\ ([\mathbf{R}_p(h_+)\mathbf{H}_p^{-1}(h_+)]'\mathbf{K}_+(h_+)\mathbf{R}_p(h_+)\mathbf{H}_p^{-1}(h_+))^{-1}[\mathbf{R}_p(h_+)\mathbf{H}_p^{-1}(h_+)]'\mathbf{K}_+(h_+)\mathbf{Y}\end{pmatrix}$$

Simplifying these expressions and denoting results using the shorthand notation yields that

$$\hat{\hat{\beta}}_{Y-,p}(h-) = \frac{1}{\sqrt{nh_-}}\mathbf{H}_p^{-1}(h_-)\mathbf{P}_{-,p}(h_-)\mathbf{Y}, \quad \hat{\hat{\beta}}_{Y+,p}(h_+) = \frac{1}{\sqrt{nh_+}}\mathbf{H}_p^{-1}(h_+)\mathbf{P}_{+,p}(h_+)\mathbf{Y}$$

Finally, under Assumptions RA-1, RA-2 and bandwidths sequences such that

$$n\min\{h_-,h_+\} \to \infty, \qquad \max\{h_-,h_+\} \to 0$$

Then

$$\hat{\hat{\beta}}_{Y-,p}(h_-) \xrightarrow{\mathbb{P}} \beta_{Y-,p}; \qquad \hat{\hat{\beta}}_{Y+,p}(h_+) \xrightarrow{\mathbb{P}} \beta_{Y+,p}$$

Which means that

$$\begin{aligned}\hat{\tau}_{Y,\nu}(\mathbf{h}) &= \nu!\,\mathbf{e}_\nu'\hat{\hat{\beta}}_{Y+,p}(\mathbf{h}) - \nu!\,\mathbf{e}_\nu'\hat{\hat{\beta}}_{Y-,p}(\mathbf{h})\\ &\xrightarrow{\mathbb{P}} \nu!\,\mathbf{e}_\nu'\beta_{Y+,p} - \nu!\,\mathbf{e}_\nu'\beta_{Y-,p}\\ &= \mu_{Y+}^{(\nu)}(0) - \mu_{Y-}^{(\nu)}(0)\end{aligned}$$

where $\mathbf{e}_\nu$ is the conformable $(\nu+1)$-th unit vector. Hence

$$\hat{\tau}_{Y,0} \xrightarrow{\mathbb{P}} \mu_{Y+}(0) - \mu_{Y-}(0)$$

B.4 Proof of Theorem 1

Proof. Calonico et al. (2019b) mention this result on page 39 without formal proof. I choose, however, to demonstrate this result because it is significant to my paper. Consider the modified covariate vector:

$$\dot{\mathbf{Z}}_d' = (X_d \times Z_{d1}, \cdots, X_d \times Z_{dm}, (1-X_d)\times Z_{d1}, \cdots, (1-X_d)\times Z_{dm})' \in \mathbb{R}^{2m}$$

The covariate-adjusted RD estimator $\hat{\hat{\theta}}_{Y,p}(\mathbf{h})$ implemented with bandwidths $\mathbf{h} = (h_-, h_+)$ that follows model (3.2) is:

$$\dot{\theta}_{Y,p}(\mathbf{h}) = \left[\begin{array}{c} \dot{\beta}_{Y,p}(\mathbf{h}) \\ \dot{\gamma}_{Y,p}(\mathbf{h}) \end{array} \right], \quad \dot{\beta}_{Y,p}(\mathbf{h}) \in \mathbb{R}^{2+2p}, \quad \dot{\gamma}_{Y,p}(\mathbf{h}) = (\dot{\gamma}_{Y,p,+}(\mathbf{h}), \dot{\gamma}_{Y,p,-}(\mathbf{h})) \in \mathbb{R}^{2m},$$

$$\hat{\dot{\theta}}_{Y,p}(\mathbf{h}) = \underset{b_-,b_+,\gamma}{argmin} \sum_{d=1}^{n} \left(Y_d - \mathbf{r}_{-,p}(\Delta_d)' \mathbf{b}_- - \mathbf{r}_{+,p}(\Delta_d)' \mathbf{b}_+ - \dot{\mathbf{Z}}_d' \gamma \right)^2 K_{\mathbf{h}}(\Delta_d),$$

where $\mathbf{b}_- \in \mathbb{R}^{1+p}, \mathbf{b}_+ \in \mathbb{R}^{1+p}, \gamma \in \mathbb{R}^{2m}$, and

$$\mathbf{r}_{-,p}(u) := \mathbb{1}(u < 0)\mathbf{r}_p(u), \quad \mathbf{r}_{+,p}(u) := \mathbb{1}(u \geq 0)\mathbf{r}_p(u).$$

so that

$$\tilde{\mathbf{R}}_{-,p}(h) = \left[\mathbb{1}(\Delta_1 < 0)\mathbf{r}_p\left(\frac{\Delta_1}{h}\right), \cdots, \mathbb{1}(\Delta_n < 0)\mathbf{r}_p\left(\frac{\Delta_n}{h}\right) \right]'_{n\times(p+1)}$$

and

$$\tilde{\mathbf{R}}_{+,p}(h) = \left[\mathbb{1}(\Delta_1 \geq 0)\mathbf{r}_p\left(\frac{\Delta_1}{h}\right), \cdots, \mathbb{1}(\Delta_n \geq 0)\mathbf{r}_p\left(\frac{\Delta_n}{h}\right) \right]'_{n\times(p+1)}$$

Let

$$\tilde{\mathbf{R}}_p(\mathbf{h}) = \begin{pmatrix} \tilde{\mathbf{R}}_{-,p}(h_-)\mathbf{H}_p^{-1}(h_-) & \tilde{\mathbf{R}}_{+,p}(h_+)\mathbf{H}_p^{-1}(h_+) \end{pmatrix}_{n\times 2(p+1)}$$

and

$$\mathbf{W}(\mathbf{h}) = (\mathbf{K}_-(h_-) + \mathbf{K}_+(h_+))_{n\times n}$$

The normal equations associated with this model are

$$\begin{pmatrix} \tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}} \\ \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}} \end{pmatrix} \begin{pmatrix} \hat{\dot{\beta}}_{Y,p}(\mathbf{h}) \\ \hat{\dot{\gamma}}_{Y,p}(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y} \\ \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{Y} \end{pmatrix}$$

where

$$\dot{\mathbf{Z}} = [X_1\mathbf{Z}_1, \cdots, X_n\mathbf{Z}_n, (1-X_1)\mathbf{Z}_1, \cdots, (1-X_n)\mathbf{Z}_n]'_{n\times 2m}$$

and

$$X_d\mathbf{Z}_d = [X_dZ_{d1}, X_dZ_{d2}, \cdots, X_dZ_{dm}]', \quad d = 1, 2, \cdots, n$$
$$(1-X_d)\mathbf{Z}_d = [(1-X_d)Z_{d1}, (1-X_d)Z_{d2}, \cdots, (1-X_d)Z_{dm}]', \quad d = 1, 2, \cdots, n$$

From these equations, we can assert that

$$\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h})\hat{\dot{\beta}}_{Y,p}(\mathbf{h})+\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\hat{\dot{\gamma}}_{Y,p}(\mathbf{h})=\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y}$$

and solving for $\hat{\dot{\beta}}_{Y,p}(\mathbf{h})$:

$$\hat{\dot{\beta}}_{Y,p}(\mathbf{h})=[\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h})]^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y}-[\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h})]^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\hat{\dot{\gamma}}_{Y,p}(\mathbf{h})$$

But notice that

$$\hat{\tilde{\beta}}_{Y,p}(\mathbf{h})=(\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}))^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y}$$

and

$$\begin{aligned}
\hat{\tilde{\beta}}'_{\dot{\mathbf{Z}},p}(\mathbf{h})&=(\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}))^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\\
\hat{\tilde{\beta}}'_{\dot{\mathbf{Z}},p}(\mathbf{h})&=\begin{pmatrix}\hat{\tilde{\beta}}_{X_dZ_{d1},p}(\mathbf{h}) & \cdots & \hat{\tilde{\beta}}_{X_dZ_{dm},p}(\mathbf{h}) & \hat{\tilde{\beta}}_{(1-X_d)Z_{d1},p}(\mathbf{h}) & \cdots & \hat{\tilde{\beta}}_{X_dZ_m,p}(\mathbf{h})\end{pmatrix}'_{(p+1)\times 2m}
\end{aligned}$$

where $\hat{\tilde{\beta}}_{Z_\ell,p}(\mathbf{h})$ is the local polynomial estimator obtained when Z_ℓ is used as the dependent variable in the method described in Appendix B.3.

Hence

$$\hat{\dot{\beta}}_{Y,p}(\mathbf{h})=\hat{\tilde{\beta}}_{Y,p}(\mathbf{h})-\hat{\tilde{\beta}}'_{\dot{\mathbf{Z}},p}(\mathbf{h})\hat{\dot{\gamma}}_{Y,p}(\mathbf{h})$$

Finally, notice that for each $l=1,\cdots,m$:

$$\begin{aligned}
\hat{\tilde{\beta}}_{X_dZ_{dl},p}(\mathbf{h})&=\hat{\tilde{\beta}}_{X_dZ_{dl},p,+}(\mathbf{h})-\hat{\tilde{\beta}}_{X_dZ_{dl},p,-}(\mathbf{h})=\hat{\tilde{\beta}}_{X_dZ_{dl},p,+}(\mathbf{h})\\
\hat{\tilde{\beta}}_{(1-X_d)Z_{dl},p}(\mathbf{h})&=\hat{\tilde{\beta}}_{(1-X_d)Z_{dl},p,+}(\mathbf{h})-\hat{\tilde{\beta}}_{(1-X_d)Z_{dl},p,-}(\mathbf{h})=-\hat{\tilde{\beta}}_{(1-X_d)Z_{dl},p,-}(\mathbf{h})
\end{aligned}$$

So that

$$\hat{\dot{\beta}}_{Y,p}(\mathbf{h})=\hat{\tilde{\beta}}_{Y,p}(\mathbf{h})-\left[\hat{\tilde{\beta}}'_{X_dZ_{dl},p,+}(\mathbf{h})\hat{\dot{\gamma}}_{Y,p,+}(\mathbf{h})-\hat{\tilde{\beta}}'_{(1-X_d)Z_{dl},p,-}(\mathbf{h})\hat{\dot{\gamma}}_{Y,p,-}(\mathbf{h})\right]$$

The result follows by examining the first component of this vector.

■

B.5 Proof of Theorem 2

I begin with a pair of lemmas:

Lemma 2.1. *Whenever model* (3.2) *is used for estimating local polynomial estimators, the following identity holds using a bandwidth sequence* $\mathbf{h} = (h_-, h_+)$*:*

$$\hat{\dot{\gamma}}_{Y,p}(\mathbf{h}) = \dot{\Gamma}_p(\mathbf{h})^{-1}\dot{\Upsilon}_{Y,p}(\mathbf{h})$$

where

$$\begin{aligned}\dot{\Gamma}_p(\mathbf{h}) =& \dot{\mathbf{Z}}'\mathbf{K}_-(h_-)\dot{\mathbf{Z}}/n - \Upsilon_{\dot{Z}-p}(h_-)'\Gamma^{-1}_{-,p}(h_-)\Upsilon_{\dot{Z}-,p}(h_-)\\ &+ \dot{\mathbf{Z}}'\mathbf{K}_+(h_+)\dot{\mathbf{Z}}/n - \Upsilon_{\dot{Z}+p}(h_+)'\Gamma^{-1}_{+,p}(h_+)\Upsilon_{\dot{Z}+,p}(h_+),\\ \dot{\Upsilon}_{Y,p}(\mathbf{h}) =& \dot{\mathbf{Z}}'\mathbf{K}_-(h_-)\mathbf{Y}/n - \Upsilon_{\dot{Z}-,p}(h_-)'\Gamma^{-1}_{-,p}(h_-)\Upsilon_{Y-p}(h_-)\\ &+ \dot{\mathbf{Z}}'\mathbf{K}_+(h_+)\mathbf{Y}/n - \Upsilon_{\dot{Z}+p}(h_+)'\Gamma^{-1}_{+,p}(h_+)\Upsilon_{Y+,p}(h_+).\end{aligned}$$

Proof. This equality is mentioned by Calonico et al. (2019a) on page 17 without formal proof. For the sake of completeness, I provide proof of this result.

Depart from the setup presented in Appendix B.4 and notice that the normal equations associated with the model are

$$\begin{pmatrix}\hat{\dot{\beta}}_{Y,p}(\mathbf{h})\\ \hat{\dot{\gamma}}_{Y,p}(\mathbf{h})\end{pmatrix} = \begin{pmatrix}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\\ \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\end{pmatrix}^{-1}\begin{pmatrix}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\mathbf{Y}\\ \mathbf{Z}'\mathbf{W}(\mathbf{h})\mathbf{Y}\end{pmatrix}$$

Following the block matrix inversion formula:

$$\begin{pmatrix}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\\ \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) & \dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}\end{pmatrix}^{-1} = \begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}^{-1} = \begin{pmatrix}A^{-1}_{11\cdot 2} & -A^{-1}_{11\cdot 2}A_{12}A^{-1}_{22}\\ -A^{-1}_{22\cdot 1}A_{21}A^{-1}_{11} & -A^{-1}_{22\cdot 1}\end{pmatrix}$$

where

$$\begin{aligned}A^{-1}_{11\cdot 2} &= (A_{11} - A_{12}A^{-1}_{22}A_{21})^{-1}\\ A^{-1}_{22\cdot 1} &= (A_{22} - A_{21}A^{-1}_{11}A_2)^{-1}\end{aligned}$$

Standard matrix calculations result in

$$\begin{pmatrix}\hat{\dot{\beta}}_{Y,p}(\mathbf{h})\\ \hat{\dot{\gamma}}_{Y,p}(\mathbf{h})\end{pmatrix} = \begin{pmatrix}[\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{M}_2\tilde{\mathbf{R}}_p(\mathbf{h})]^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{M}_2\mathbf{Y}\\ [\dot{\mathbf{Z}}'\mathbf{M}_1\dot{\mathbf{Z}}]^{-1}\dot{\mathbf{Z}}'\mathbf{M}_1\mathbf{Y}\end{pmatrix}$$

where

$$\mathbf{M}_1 = \mathbf{W}(\mathbf{h}) - \mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h})(\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}))^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})$$
$$\mathbf{M}_2 = \mathbf{W}(\mathbf{h}) - \mathbf{W}(\mathbf{h})\dot{\mathbf{Z}}(\dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\dot{\mathbf{Z}})^{-1}\dot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})$$

Now, using the identity

$$\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}) = \mathbf{K}_-(h_-)\mathbf{R}_p(h_-)\mathbf{H}_p^{-1}(h_-) + \mathbf{K}_+(h_+)\mathbf{R}_p(h_+)\mathbf{H}_p^{-1}(h_+)$$

then standard simplifications yield

$$\hat{\dot{\gamma}}_{Y,p}(\mathbf{h}) = [\dot{\mathbf{Z}}'\mathbf{M}_1\mathbf{Z}]^{-1}\dot{\mathbf{Z}}'\mathbf{M}_1\mathbf{Y} = \dot{\Gamma}_p(\mathbf{h})^{-1}\dot{\Upsilon}_{Y,p}(\mathbf{h})$$

■

It follows from Lemma 2.1 that

$$\hat{\dot{\gamma}}_{Y,p,+}(\mathbf{h}) = \dot{\Gamma}_{p+}(\mathbf{h})^{-1}\dot{\Upsilon}_{Y,p,+}(\mathbf{h})$$
$$\hat{\dot{\gamma}}_{Y,p,-}(\mathbf{h}) = \dot{\Gamma}_{p-}(\mathbf{h})^{-1}\dot{\Upsilon}_{Y,p,-}(\mathbf{h})$$

where

$$\dot{\Gamma}_{p,-}(\mathbf{h}) = \mathbf{Z}'\mathbf{K}_-(h_-)\mathbf{Z}/n - \Upsilon_{Z-p}(h_-)'\Gamma_{-,p}^{-1}(h_-)\Upsilon_{Z-,p}(h_-)$$
$$\dot{\Gamma}_{p,+}(\mathbf{h}) = \mathbf{Z}'\mathbf{K}_+(h_+)\mathbf{Z}/n - \Upsilon_{Z+p}(h_+)'\Gamma_{+,p}^{-1}(h_+)\Upsilon_{Z+,p}(h_+),$$
$$\dot{\Upsilon}_{Y,p,-}(\mathbf{h}) = \mathbf{Z}'\mathbf{K}_-(h_-)\mathbf{Y}/n - \Upsilon_{Z-,p}(h_-)'\Gamma_{-,p}^{-1}(h_-)\Upsilon_{Y-p}(h_-)$$
$$\dot{\Upsilon}_{Y,p,+}(\mathbf{h}) = \mathbf{Z}'\mathbf{K}_+(h_+)\mathbf{Y}/n - \Upsilon_{Z+p}(h_+)'\Gamma_{+,p}^{-1}(h_+)\Upsilon_{Y+,p}(h_+).$$

The following lemma characterizes these estimators asymptotically.

Lemma 2.2. *Under regularity assumptions RA-1, RA-2, RA-3 and bandwidth sequences* $\mathbf{h}$ *such that* $n\min\{h_-,h_+\} \to \infty$ *and* $\max\{h_-,h_+\} \to 0$

$$\dot{\Gamma}_{p,+}(\mathbf{h}) = a\sigma_{Z+}^2 + o_{\mathbb{P}}(1)$$
$$\dot{\Gamma}_{p,-}(\mathbf{h}) = a\sigma_{Z-}^2 + o_{\mathbb{P}}(1)$$

and

$$\dot{\Upsilon}_{Y,p,+}(\mathbf{h}) = a\mathbb{E}\left[\left(\mathbf{Z}_d(0) - \mu_{Z-}(\Delta_d)\right)Y_d(0)|\Delta_d = 0\right] + o_{\mathbb{P}}(1)$$
$$\dot{\Upsilon}_{Y,p,-}(\mathbf{h}) = a\mathbb{E}\left[\left(\mathbf{Z}_d(1) - \mu_{Z+}(\Delta_d)\right)Y_d(1)|\Delta_d = 0\right] + o_{\mathbb{P}}(1)$$

where

$$a = f(0)\int_{-\infty}^{0} K(u)du = f(0)\int_{0}^{\infty} K(u)du$$

and

$$\mu_{Z-} = \mu_{Z-}(0), \quad \mu_{Z+} = \mu_{Z+}(0), \quad \sigma^2_{Z-} = \sigma^2_{Z-}(0), \quad \sigma^2_{Z+} = \sigma^2_{Z+}(0)$$

Proof. See Lemma SA-6 of Calonico et al. (2019a). ∎

I now proceed with the Theorem's proof:

Proof. From Lemmas 2.1 and 2.2 it follows that

$$\begin{aligned}
\hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}) &\xrightarrow{\mathbb{P}} (\sigma^2_{Z-})^{-1}\mathbb{E}\left[\left(\mathbf{Z}_d(0) - \mu_{Z-}(\Delta_d)\right)Y_d(0)|\Delta_d = 0\right] \\
\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) &\xrightarrow{\mathbb{P}} (\sigma^2_{Z+})^{-1}\mathbb{E}\left[\left(\mathbf{Z}_d(1) - \mu_{Z+}(\Delta_d)\right)Y_d(1)|\Delta_d = 0\right]
\end{aligned}$$

I only prove the sub-index "+" case since the other case is identical. From Assumption $2'$ it can be deduced that

$$\begin{aligned}
\mathbb{E}\left[\left(\mathbf{Z}_d(1) - \mu_{Z+}(\Delta_d)\right)Y_d(1)|\Delta_d = 0\right] &= \mathbb{E}\left[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)Y_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] \\
&= \mathbb{E}\left[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)(\tau_d + \mathbf{Z}'_{id}\gamma)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] \\
&= \mathbb{E}\left[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)\mathbf{Z}'_{id}\gamma)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] \\
&= \mathbb{E}\Big[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)'\gamma)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\Big] \\
&+ \mathbb{E}\left[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)\mu_{Z+}(\Delta_{id})'\gamma)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\right] \\
&= \mathbb{E}\Big[\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)\left(\mathbf{Z}_{id} - \mu_{Z+}(\Delta_{id})\right)'|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0\Big]\gamma \\
&= \mathbb{V}[\mathbf{Z}_{id}(1)|\Delta_{id} = 0, X_{id} = 1, X_{jd} = 0]\gamma \\
&= \mathbb{V}[\mathbf{Z}_d(1)|\Delta_d = 0]\gamma \\
&= \sigma^2_{Z+}\,\gamma
\end{aligned}$$

Therefore

$$\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) \xrightarrow{\mathbb{P}} (\sigma^2_{Z+})^{-1}\sigma^2_{Z+}\gamma = \gamma$$

∎

B.6 Proof of Theorem 3

Proof. From Theorem 1, we know that

$$\hat{\hat{\tau}}_Y = \hat{\tau}_{PCRD} - \left[(\hat{\hat{\tau}}_{\mathbf{Z},+})' \hat{\hat{\gamma}}_+ - (\hat{\hat{\tau}}_{\mathbf{Z},-})' \hat{\hat{\gamma}}_- \right]$$

and from Theorem 2, we have established that

$$\hat{\hat{\gamma}}(\mathbf{h}) \xrightarrow{\mathbb{P}} \gamma$$

Moreover, making use of Assumption 3

$$\hat{\hat{\tau}}'_{\mathbf{Z},+} \xrightarrow{\mathbb{P}} \delta'_+, \quad \hat{\hat{\tau}}'_{\mathbf{Z},+} \xrightarrow{\mathbb{P}} \delta'_-$$

Hence

$$(\hat{\hat{\tau}}_{\mathbf{Z},+})' \hat{\hat{\gamma}}_+ - (\hat{\hat{\tau}}_{\mathbf{Z},-})' \hat{\hat{\gamma}}_- \xrightarrow{\mathbb{P}} \delta'_+ \gamma - \delta'_- \gamma = (\delta_+ - \delta_-)' \gamma = \delta' \gamma = b_{PCRD}$$

Using Assumption 1, this means that

$$\hat{\hat{\tau}}_Y \xrightarrow{\mathbb{P}} (\tau_{PCRD} + b_{PCRD}) - b_{PCRD} = \tau_{PCRD}$$

■

B.7 Proof of Theorem 4

Proof. I provide an alternative proof to the one provided in Frölich and Huber (2019b, p.2), using the alternative set of regularity assumptions listed in RA-4 and Assumption 4.

Let $U_\eta = [-\eta, \eta]$ be a symmetric closed neighborhood around 0 of length $\eta > 0$. Further denote by $U_\eta^+ = [0, \eta]$ and $U_\eta^- = [-\eta, 0]$. Using assumption *a*) of RA-4

$$\begin{aligned}
\lim_{\eta \to 0} \mathbb{P}(\Delta_d \geq 0 | \Delta_d \in U_\eta) &= \lim_{\eta \to 0} \frac{F_\Delta(\eta) - F_\Delta(0)}{F_\Delta(\eta) - F_\Delta(-\eta)} \\
&= \frac{[F_\Delta(\eta) - F_\Delta(0)]/\eta}{[F_\Delta(\eta) - F_\Delta(0)]/\eta - [F_\Delta(-\eta) - F_\Delta(0)]/\eta} \\
&= \frac{f_\Delta(0)}{f_\Delta(0) + f_\Delta(0)} \\
&= \frac{1}{2}
\end{aligned}$$

Hence, leveraging assumption *b*) of RA-4

$$
\begin{aligned}
\lim_{\eta\to 0}\mathbb{P}(\mathbf{Z}_d\leq z|\Delta_d\in U_\eta) &= \lim_{\eta\to 0}\mathbb{P}(\mathbf{Z}_d\leq z|\Delta_d\in U_\eta^-)\mathbb{P}(\Delta_d\leq 0|\Delta_d\in U_\eta)\\
&+\lim_{\eta\to 0}\mathbb{P}(\mathbf{Z}_d\leq z|\Delta_d\in U_\eta^+)\mathbb{P}(\Delta_d> 0|\Delta_d\in U_\eta)\\
&=\frac{F^-_{\mathbf{Z}_d(0)|\Delta_d}(z|0)+F^+_{\mathbf{Z}_d(1)|\Delta_d}(z|0)}{2}
\end{aligned}
$$

Following assumption $c)$ of RA-4, then

$$
\lim_{\eta\to 0} f_{\mathbf{Z}|\Delta_d\in U_\eta}(z)=\frac{f^+_{\mathbf{Z}_d(0)|\Delta_d}(z|0)-f^+_{\mathbf{Z}_d(1)|\Delta_d}(z|0)}{2}
$$

Next, given assumption $d)$ of RA-4, the dominated convergence theorem allows to switch the limit and integral operators. Hence, invoking Assumption 4

$$
\begin{aligned}
&\lim_{\eta\to 0}\int\left(\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^+]-\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^-]\right)f_{\mathbf{Z}|\Delta_d\in U_\eta}(z)dz=\\
&\int\lim_{\eta\to 0}\left(\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^+]-\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^-]\right)\lim_{\eta\to 0}f_{\mathbf{Z}|\Delta_d\in U_\eta}(z)dz=\\
&\int\left(m^+(z,0)-m^-(z,0)\right)\left(\frac{f^-_{\mathbf{Z}_d(0)|\Delta_d}(z|0)+f^+_{\mathbf{Z}_d(1)|\Delta_d}(z|0)}{2}\right)dz
\end{aligned}
$$

Likewise, observe that

$$
\mathbb{E}[Y_d(1)-Y_d(0)|\Delta_d=0]=\int\mathbb{E}[Y_d(1)-Y_d(0)|\Delta_d=0,\mathbf{Z}_d=z]f_{\mathbf{Z}|\Delta_d=0}(z)dz
$$

Thus,

$$
\begin{aligned}
&\int\left(\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^+]-\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^-]\right)f_{\mathbf{Z}|\Delta_d\in U_\eta}(z)dz=\\
&\int\left(\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^+]-\mathbb{E}[Y_d|\mathbf{Z}_d=z,\Delta_d\in U_\eta^-]\right)f_{\mathbf{Z}|\Delta_d\in U_\eta}(z)dz+\\
&\int\mathbb{E}[Y_d(1)-Y_d(0)|\Delta_d=0,\mathbf{Z}_d=z]f_{\mathbf{Z}|\Delta_d=0}(z)dz-\\
&\int\mathbb{E}[Y_d(1)-Y_d(0)|\Delta_d=0,\mathbf{Z}_d=z]f_{\mathbf{Z}|\Delta_d=0}(z)dz
\end{aligned}
$$

where the third and fourth lines add to zero. Rearranging terms leads to

$$\int \left(\mathbb{E}[Y_d|\mathbf{Z}_d = z, \Delta_d \in U_\eta^+] - \mathbb{E}[Y_d|\mathbf{Z}_d = z, \Delta_d \in U_\eta^-]\right) f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) dz =$$
$$\mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d \in U_\eta] +$$
$$\int \left(\mathbb{E}[Y_d(1)|\mathbf{Z}_d = z, \Delta_d \in U_\eta^+] f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) - \mathbb{E}[Y_d(1)|\mathbf{Z}_d = z, \Delta_d = 0] f_{\mathbf{Z}|\Delta_d = 0}(z)\right) dz -$$
$$\int \left(\mathbb{E}[Y_d(0)|\mathbf{Z}_d = z, \Delta_d \in U_\eta^-] f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) - \mathbb{E}[Y_d(0)|\mathbf{Z}_d = z, \Delta_d = 0] f_{\mathbf{Z}|\Delta_d = 0}(z)\right) dz$$

Using the dominated convergence theorem again and Assumption 4

$$\lim_{\eta \to 0} \int \left(\mathbb{E}[Y_d(1)|\mathbf{Z}_d = z, \Delta_d \in U_\eta^+] f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) - \mathbb{E}[Y_d(1)|\mathbf{Z}_d = z, \Delta_d = 0] f_{\mathbf{Z}|\Delta_d = 0}(z)\right) dz = 0$$

and

$$\lim_{\eta \to 0} \int \left(\mathbb{E}[Y_d(0)|\mathbf{Z}_d = z, \Delta_d \in U_\eta^-] f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) - \mathbb{E}[Y_d(0)|\mathbf{Z}_d = z, \Delta_d = 0] f_{\mathbf{Z}|\Delta_d = 0}(z)\right) dz = 0$$

Consequently,

$$\lim_{\eta \to 0} \int \left(\mathbb{E}[Y_d|\mathbf{Z}_d = z, \Delta_d \in U_\eta^+] - \mathbb{E}[Y_d|\mathbf{Z}_d = z, \Delta_d \in U_\eta^-]\right) f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) dz = \mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0]$$

which completes the proof in that

$$\mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0] = \int \left(m^+(z,0) - m^-(z,0)\right) \left(\frac{f^-_{\mathbf{Z}_d(0)|\Delta_d}(z|0) + f^+_{\mathbf{Z}_d(1)|\Delta_d}(z|0)}{2}\right) dz$$

■

B.8 Proof of Theorem 5

Proof. While it is possible to demonstrate this assertion directly, I will opt to prove that the suggested estimator is asymptotically equivalent to the estimators proposed by Frölich and Huber (2019a), which are consistent estimators of $\mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0]$ whenever Assumption 4 holds. This serves two purposes. First, the authors already have detailed proof in their article relating the asymptotic properties of their estimators. Second, illustrating this relationship helps to understand better why the estimator works.

To begin with, Frölich and Huber (2019a) study non-parametric estimators that can potentially estimate the unconditional causal effect $\mathbb{E}[Y_d(1) - Y_d(0)|\Delta_d = 0]$ following the result illustrated in Theorem 4. In particular, they find that provided some regularity assumptions, all estimators $\hat{\theta}$ of the class of estimators described in (B.1) can consistently estimate this effect.

$$\hat{\theta} = \frac{\sum_{d=1}^{n} \left(\hat{m}^{+}(\mathbf{Z}_d, 0) - \hat{m}^{-}(\mathbf{Z}_d, 0)\right) \bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}{\sum_{d=1}^{n} \bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)} \tag{B.1}$$

These estimators are obtained from a two-step procedure that mimics the logic induced by Theorem 4. First, the researcher must estimate non-parametrically the functions $m^{+}(\cdot,0)$ and $m^{-}(\cdot,0)$ using a non-parametric regression method. Next, the resulting estimates $\hat{m}^{+}(\cdot,0)$ and $\hat{m}^{-}(,0)$ are evaluated at the observed values of the covariates and averaged over using a kernel function $\bar{K}(\cdot)$ over a bandwidth $\bar{h} > 0$.

In particular, Frölich and Huber (2019a) suggests estimating $m^{+}(\cdot,0)$ using local polynomials. For any fixed z, $\hat{m}^{+}(z,0)$ is the value $a(z)$ that solves

$$\underset{a(z),b(z),c(z)}{argmin} \sum_{d=1}^{n} \left(Y_d - a(z) - b(z)\Delta_d - (\mathbf{Z}_d - z)'c(z)\right)^2 K_d \mathbb{1}[\Delta_d \geq 0]$$

where K_d is a product kernel

$$K_d = K_h(\Delta_d) \prod_{l=1}^{m} \tilde{K}_{h_z}\left(\frac{Z_{dl} - z_l}{h_z}\right)$$

and $\tilde{K}_{h_z}(\cdot)$ is a kernel of order greater than 2. The estimate for $m^{-}(\cdot,0)$ is determined in the same way, using the sample at the left side of the threshold.

I summarize the consistency results for estimators of class (B.1) in the following lemma.

Lemma 5.1 (Frölich and Huber (2019a), Proposition 2.)**.** *Under Assumption 4, and regularity conditions RA-4 and RA-5, estimators of the class* (B.1) *are consistent estimators of* τ_{PCRD}*.*

Proof. I refer to the original proof in the article. However, the intuition follows from the proof of Theorem 4. Notice that, for an interval $U_\eta = [-\eta, \eta]$, $\eta > 0$:

$$\begin{aligned}\lim_{\bar{h}\to 0} \frac{\sum_{d=1}^{n} \left(\hat{m}^{+}(\mathbf{Z}_d, 0) - \hat{m}^{-}(\mathbf{Z}_d, 0)\right) \bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}{\sum_{d=1}^{n} \bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)} &\approx \lim_{\eta\to 0} \int \left(\hat{m}^{+}(z,0) - \hat{m}^{-}(z,0)\right) f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) dz \\ &\approx \lim_{\eta\to 0} \int \left(m^{+}(z,0) - m^{-}(z,0)\right) f_{\mathbf{Z}|\Delta_d \in U_\eta}(z) dz = \tau_{PCRD}\end{aligned}$$

■

Now consider the following result for linear-in-parameters covariate-adjusted polynomial estimators.

Lemma 5.2. *For any vector $z \in \mathbb{R}^m$, the algebraic solution to the problem*

$$\underset{b_-,b_+,\gamma}{argmin} \sum_{d=1}^{n} \left(Y_d - \mathbf{r}_{-,p}(\Delta_d)' \mathbf{b}_- - \mathbf{r}_{+,p}(\Delta_d)' \mathbf{b}_+ - (\dot{\mathbf{Z}}_d - \dot{z})'\gamma \right)^2 K_{\mathbf{h}}(\Delta_d)$$

is

$$\hat{\hat{\beta}}_{Y,p}(\mathbf{h})[z] = \hat{\hat{\beta}}_{Y,p}(\mathbf{h}) - \left[\hat{\hat{\beta}}'_{X_d Z_{dl},p,+}(\mathbf{h}) \hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\beta}}'_{(1-X_d) Z_{dl},p,-}(\mathbf{h}) \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}) \right] + z'(\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}))$$

where $\dot{z} = (z,z)'$.

Proof. This result follows directly from the procedure described in Appendix B.4, by changing $\mathbf{Z}_{id}$ to $\mathbf{Z}_{id} - z$.

■

Lemma 5.2 is enough to prove part a). I now describe the link between these estimators and a particular linear-in-parameters specification. Let

$$\bar{\mathbf{Z}}_{NW} = \left(\sum_{d=1}^{n} \mathbf{Z}_d \bar{K}\left(\frac{\Delta_d - 0}{\bar{h}}\right) \right) \Big/ \sum_{d=1}^{n} \bar{K}\left(\frac{\Delta_d - 0}{\bar{h}}\right),$$

be an estimate of the average characteristics at the threshold using a Nadaraya–Watson kernel estimator. This procedure consists in averaging characteristics using a kernel function $\bar{K}(\cdot)$ over a bandwidth $\bar{h} > 0$, aiming to "localize" the average around the point $\Delta_d = 0$.

Lemma 5.3. *The estimator* $\hat{\hat{\tau}}_{Y,p}(\mathbf{h})[\bar{\mathbf{Z}}_{NW}]$ *belongs to the class* (B.1).

Proof. Since local polynomial estimators consistently estimate the regression functions under Regularity assumptions RA-4 and RA-5, it follows that

$$\hat{m}^+(z,0) = \mathbf{e}_0' \left[\hat{\hat{\beta}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\beta}}'_{X_d Z_{dl},p,+}(\mathbf{h}) \hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) + z' \hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) \right] \xrightarrow{\mathbb{P}} m^+(z,0)$$

and

$$\hat{m}^-(z,0) = \mathbf{e}_0' \left[\hat{\hat{\beta}}_{Y,p,-}(\mathbf{h}) - \hat{\hat{\beta}}'_{X_d Z_{dl},p,-}(\mathbf{h}) \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}) + z' \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}) \right] \xrightarrow{\mathbb{P}} m^-(z,0)$$

Hence

$$\hat{\hat{\tau}}_{Y,p}(\mathbf{h})[z] = \mathbf{e}_0' \hat{\hat{\beta}}_{Y,p}(\mathbf{h})[z] = \hat{m}^+(z,0) - \hat{m}^-(z,0)$$

Following Lemma 5.2, the estimator resulting from the local model (3.4) is

$$\hat{\hat{\tau}}_{Y,p}(\mathbf{h})[\bar{\mathbf{Z}}_{NW}] = \hat{\hat{\tau}}_{Y,p}(\mathbf{h}) - \left[\hat{\hat{\tau}}'_{X_d Z_{dl},p,+}(\mathbf{h})\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\tau}}'_{(1-X_d)Z_{dl},p,-}(\mathbf{h})\hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h})\right]$$
$$+ \bar{\mathbf{Z}}'_{NW}(\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}))$$

However, using the definition of $\bar{\mathbf{Z}}_{NW}$, the estimator can be reformulated as

$$\hat{\hat{\tau}}_{Y,p}(\mathbf{h})[\bar{\mathbf{Z}}_{NW}]$$
$$= \frac{\sum_{d=1}^{n}\left[\hat{\hat{\tau}}_{Y,p}(\mathbf{h}) - \left[\hat{\hat{\tau}}'_{X_d Z_{dl},p,+}(\mathbf{h})\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\tau}}'_{(1-X_d)Z_{dl},p,-}(\mathbf{h})\hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h})\right] + \mathbf{Z}'_d(\hat{\hat{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\hat{\gamma}}_{Y,p,-}(\mathbf{h}))\right]\bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}{\sum_{d=1}^{n}\bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}$$
$$= \frac{\sum_{d=1}^{n}(\hat{m}^{+}(\mathbf{Z}_d,0) - \hat{m}^{-}(\mathbf{Z}_d,0))\bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}{\sum_{d=1}^{n}\bar{K}\left(\frac{\Delta_d}{\bar{h}}\right)}$$

Consequently, we have proven that $\hat{\hat{\tau}}_Y(\bar{\mathbf{Z}}_{NW})$ belongs to the estimator class (B.1)[21].

■

To end the argument, if $\bar{h} = h$ and $\bar{K}(\cdot) = K(\cdot)$, note that $\bar{\mathbf{Z}}_{NW}$ can be written using local polynomial estimators of degree $p = 0$

$$\bar{\mathbf{Z}}_{NW} = \frac{\hat{\hat{\tau}}_{X_d\mathbf{Z},0,+}(h,h) + \hat{\hat{\tau}}_{(1-X_d)\mathbf{Z},0,-}(h,h)}{2}$$

But this yields the same limit in probability that using a local polynomial estimator of degree p. Hence

$$\bar{\mathbf{Z}} = \frac{\hat{\hat{\tau}}_{X_d\mathbf{Z},p,+}(\mathbf{h}) + \hat{\hat{\tau}}_{(1-X_d)\mathbf{Z},p,-}(\mathbf{h})}{2} = \bar{\mathbf{Z}}_{NW} + o_{\mathbb{P}}(1)$$

so that, a direct application of the continuous mapping theorem results in

$$\hat{\tau}^{PC}_{PCRD} = \hat{\hat{\tau}}_{Y,p}(\mathbf{h})[\bar{\mathbf{Z}}] = \hat{\hat{\tau}}_{Y,p}(\mathbf{h})[\bar{\mathbf{Z}}_{NW}] + o_{\mathbb{P}}(1) = \tau_{PCRD} + o_{\mathbb{P}}(1)$$

This proves part *b*). Part *c*) is an immediate consequence of parts *a*) and *b*).

■

B.9 Bandwidth choice and asymptotic normality of the bias-corrected estimator

Since the asymptotics of covariate-adjusted local polynomial estimators has been extensively developed in Calonico et al. (2019b), the bias-corrected estimator's specific structure makes inference fairly straightfor-

[21] $\bar{K}(\cdot)$ is the uniform kernel in this application.

ward. This Appendix briefly shows how their results can be adapted to my work.

The first observation is that all relevant estimators share a common functional form. To see this, recall from Theorem 5 that

$$\hat{\tau}^{BC}_{PCRD} = \hat{\hat{\tau}}_Y(\mathbf{h}) - \frac{1}{2}\left[(\hat{\hat{\tau}}_{\mathbf{Z},+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z},-}(\mathbf{h}))'(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h}))\right]$$

Which in vector terms can be expressed as

$$\hat{\tau}^{BC}_{PCRD} = \begin{pmatrix} 1 & -\frac{1}{2}(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h})) \end{pmatrix} \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix} = \hat{\mathbf{t}}_1'(\mathbf{h}) \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix}$$

Similarly, the bias estimate of the bias

$$\begin{aligned} \hat{b}_{PCRD} &= \frac{1}{2}\left[(\hat{\hat{\tau}}_{\mathbf{Z},+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z},-}(\mathbf{h}))'(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h}))\right] \\ &= \begin{pmatrix} 0 & \frac{1}{2}(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h})) \end{pmatrix} \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix} = \hat{\mathbf{t}}_2'(\mathbf{h}) \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix} \end{aligned}$$

Finally, the bias of a particular component of the bias $\hat{b}_l$ can be expressed as

$$\hat{b}_l = \begin{pmatrix} 0 & 0 & \cdots & \frac{1}{2}[(\hat{\hat{\gamma}}_-)_l + (\hat{\hat{\gamma}}_+)_l] & \cdots & 0 \end{pmatrix} \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix} = \hat{\mathbf{t}}_{3,l}'(\mathbf{h}) \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix}$$

As a result, in general terms, all the estimators derived can are member of the family

$$\mathcal{F}_{p,m} = \left\{ \hat{\theta} : \hat{\theta} = \hat{\mathbf{t}}_{\hat{\theta}} \begin{pmatrix} \hat{\hat{\tau}}_Y(\mathbf{h}) \\ \hat{\hat{\tau}}_{\mathbf{Z}}(\mathbf{h}) \end{pmatrix}, \quad \hat{\mathbf{t}}_{\hat{\theta}} \xrightarrow{\mathbb{P}} \mathbf{t}_{\hat{\theta}} \in \mathbb{R}^{m+1} \right\}$$

where p indicates the polynomial degree used for the local polynomial estimators and m is the number of covariates. Finally, for notation consistency, denote $\theta = \underset{n\to\infty}{\text{p-lim}}\, \hat{\theta}$.

B.9.1 Bandwidth choice

Bandwidth sequence choices for local polynomial estimators are usually based on some optimality criterion. The most common target is to minimize the Mean Square Error (MSE) to make the best out of the bias-variance tradeoff these estimators face (Calonico et al., 2014; Imbens & Kalyanaraman, 2012). In brief, as bandwidth sizes shrink, the local model better approximates the conditional expectation function at the threshold, thus reducing bias. However, as the estimation windows narrow, fewer observations available increase the estimator's variability. The MSE-optimal estimator is then the one that balances most effectively these countervailing forces.

Theorem A-1 (Theorem 1, Calonico et al. (2019b))**.** *Suppose that regularity assumptions RA-1, RA-2, and RA-2 hold. Then for any $\hat{\theta} \in \mathcal{F}_{p,m}$ and bandwidth $h = h_- = h_+$.*

$$MSE[\hat{\theta}(h)] = h^{2(1+p)}\mathcal{B}_{\hat{\theta}}^2[1+o_{\mathbb{P}}(1)] + \frac{1}{nh}\mathcal{V}_{\hat{\theta}}$$

where $\mathcal{B}_{\hat{\theta}}$ and $\mathcal{V}_{\hat{\theta}}$ are the pre-asymptotic stochastic approximations to the conditional bias and variances respectively.

Theorem A-1 implies that the optimal bandwidth satisfies

$$h_{MSE}^* = \left[\frac{1}{2n(1+p)}\frac{\mathcal{V}_{\hat{\theta}}}{\mathcal{B}_{\hat{\theta}}^2[1+o_{\mathbb{P}}(1)]}\right]^{\frac{1}{3+2p}}$$

This means that a plug-in estimator can be formulated using consistent estimates $\hat{\mathcal{B}}$ and $\hat{\mathcal{V}}$ of $\mathcal{V}_{\hat{\theta}}$ and $\mathcal{B}_{\hat{\theta}}$ through the formula

$$\hat{h}_{MSE}^* = \left[\frac{1}{2n(1+p)}\frac{\hat{\mathcal{V}}}{\hat{\mathcal{B}}^2}\right]^{\frac{1}{3+2p}}$$

For the case in which bandwidths can differ at the left and right of the threshold, note that $\mathcal{B}_{\hat{\theta}} = \mathcal{B}_{\hat{\theta}_+} - \mathcal{B}_{\hat{\theta}_-}$, this is the bias is the difference of the bias observed at each side of the threshold. Similarly, $\mathcal{V}_{\hat{\theta}} = \mathcal{V}_{\hat{\theta}_+} + \mathcal{V}_{\hat{\theta}_-}$. So different bandwidths can be formulated using appropriate estimators as follows:

$$\hat{h}_{-,MSE}^* = \left[\frac{1}{2n(1+p)}\frac{\hat{\mathcal{V}}_-}{\hat{\mathcal{B}}_-^2}\right]^{\frac{1}{3+2p}}, \qquad \hat{h}_{+,MSE}^* = \left[\frac{1}{2n(1+p)}\frac{\hat{\mathcal{V}}_+}{\hat{\mathcal{B}}_+^2}\right]^{\frac{1}{3+2p}}$$

The exact formulas for these estimators will be described in the next subsection.

B.9.2 Asymptotic normality

To develop valid asymptotic distributional approximations and inference procedures, Calonico et al. (2019b) rely on nonparametric robust bias correction as suggested by Calonico et al. (2014) and Calonico, Cattaneo, and Farrell (2018). This adjustment is necessary because MSE-optimal bandwidths are systematically too "large", leading to a first-order bias in the distributional approximation. Therefore, this bias must be estimated and removed for the inference to be valid.

Consider the general formulation proved in Appendix B.4 of the equivalence formula formulated in Theorem 1:

$$\hat{\tilde{\beta}}_{Y,p}(\mathbf{h}) = \hat{\tilde{\beta}}_{Y,p}(\mathbf{h}) - \left[\hat{\tilde{\beta}}'_{X_dZ_{dl},p,+}(\mathbf{h})\hat{\tilde{\gamma}}_{Y,p,+}(\mathbf{h}) - \hat{\tilde{\beta}}'_{(1-X_d)Z_{dl},p,-}(\mathbf{h})\hat{\tilde{\gamma}}_{Y,p,-}(\mathbf{h})\right]$$

This means that

$$\hat{\tau}_{PCRD}^{BC} = \mathbf{T}_1(\mathbf{h})\mathrm{vec}(\hat{\tilde{\beta}}_{S,p}(\mathbf{h}))$$

where

$$\mathbf{T}_1(\mathbf{h}) = \begin{pmatrix} \mathbf{e}_0 \\ -\frac{1}{2}(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h})) \otimes \mathbf{e}_0 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{2}(\hat{\hat{\gamma}}_+(\mathbf{h}) + \hat{\hat{\gamma}}_-(\mathbf{h})) \end{pmatrix} \otimes \mathbf{e}_0 = (\hat{\mathbf{t}}_1(\mathbf{h}) \otimes \mathbf{e}_0)$$

$\otimes$ denotes the Kronecker product, and

$$\hat{\hat{\beta}}_{S,p}(\mathbf{h}) = \hat{\hat{\beta}}_{S,+,p}(h_+) - \hat{\hat{\beta}}_{S,-,p}(h_-)$$

with

$$\hat{\hat{\beta}}_{S,+,p}(h) = [\hat{\hat{\beta}}_{Y,+,p}(h), \hat{\hat{\beta}}_{\mathbf{Z},+,p}(h)], \qquad \hat{\hat{\beta}}_{S,-,p}(h) = [\hat{\hat{\beta}}_{Y,-,p}(h), \hat{\hat{\beta}}_{\mathbf{Z},-,p}(h)]$$

Following this exposition, any estimator $\hat{\theta} \in \mathcal{F}_{p,m}$ can be expressed as

$$\hat{\theta}(\mathbf{h}) = (\mathbf{t}_{\hat{\theta}} \otimes \mathbf{e}_0)\mathrm{vec}(\hat{\hat{\beta}}_{S,p}(\mathbf{h}))$$

Now, given the matrix expressions of the local polynomials described in Appendix B.3, it directly follows from the identity $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B)$ (for conformable matrices A, B, C) that

$$\mathrm{vec}(\hat{\hat{\beta}}_{S,-,p}(h)) = \frac{1}{\sqrt{nh}}\left[\mathbf{I}_{1+d} \otimes \mathbf{H}_p^{-1}(h)\mathbf{P}_{-,p}(h)\right]\mathbf{S},$$
$$\mathrm{vec}(\hat{\hat{\beta}}_{S,+,p}(h)) = \frac{1}{\sqrt{nh}}\left[\mathbf{I}_{1+d} \otimes \mathbf{H}_p^{-1}(h)\mathbf{P}_{+,p}(h)\right]\mathbf{S}$$

with $\mathbf{I}_d$ denoting the identity matrix of dimension d.

I know present a series of results that characterize estimators $\hat{\theta} \in \mathcal{F}_{p,m}$ by slightly modifying their original formulations presented in Calonico et al. (2019a):

The pre-asymptotic bias $\mathcal{B}_{\hat{\theta},p}(h) = \mathcal{B}_{\hat{\theta}_+,p}(h) - \mathcal{B}_{\hat{\theta}_-,p}(h)$ and its asymptotic counterpart $\mathcal{B}_{\hat{\theta},p} := \mathcal{B}_{\hat{\theta}_+,p} - \mathcal{B}_{\hat{\theta}_-,p}$ are characterized by

$$\mathcal{B}_{\hat{\theta}_-,p}(h) := \mathbf{e}_0'\Gamma_{-,p}^{-1}(h)\vartheta_{-,p}(h)\frac{\mathbf{t}_{\hat{\theta}}'\mu_{S-}^{(p+1)}}{(p+1)!} \xrightarrow{\mathbb{P}} \mathcal{B}_{\hat{\theta}_-,p},$$
$$\mathcal{B}_{\hat{\theta}_+,p}(h) := \mathbf{e}_0'\Gamma_{+,p}^{-1}(h)\vartheta_{+,p}(h)\frac{\mathbf{t}_{\hat{\theta}}'\mu_{S+}^{(p+1)}}{(p+1)!} \xrightarrow{\mathbb{P}} \mathcal{B}_{\hat{\theta}_+,p}$$

where $\mu_{S-} = \mu_{S-}(0)$ and $\mu_{S+} = \mu_{S+}(0)$.

Similarly, the pre-asymptotic variance $\mathcal{V}_{\hat{\theta},p}(h) = \mathcal{V}_{\hat{\theta}_-,p}(h) + \mathcal{V}_{\hat{\theta}_+,p}(h)$ and its asymptotic counterpart $\mathcal{V}_{\hat{\theta},p} := \mathcal{V}_{\hat{\theta}_-,p} + \mathcal{V}_{\hat{\theta}_+,p}$, are characterized by

$$\mathcal{V}_{\hat{\theta}_-,p}(h) := \left[\mathbf{t}_{\hat{\theta}}' \otimes \mathbf{e}_0'\mathbf{P}_{-,p}(h)\right]\Sigma_{S-}\left[\mathbf{t}_{\hat{\theta}} \otimes \mathbf{P}_{-,p}(h)\mathbf{e}_0\right] \xrightarrow{\mathbb{P}} \mathcal{V}_{\hat{\theta}_-,p},$$
$$\mathcal{V}_{\hat{\theta}_+,p}(h) := \left[\mathbf{t}_{\hat{\theta}}' \otimes \mathbf{e}_0'\mathbf{P}_{+,p}(h)\right]\Sigma_{S_+}\left[\mathbf{t}_{\hat{\theta}} \otimes \mathbf{P}_{+,p}(h)\mathbf{e}_0\right] \xrightarrow{\mathbb{P}} \mathcal{V}_{\hat{\theta}_+,p}$$

Notice that pre-asymptotic variances can be estimated from the data, so that $\hat{\mathcal{V}}_{\hat{\theta}}(h) = \mathcal{V}_{\hat{\theta}}(h)$. On the contrary, pre-asymptotic biases involve unknown quantities. Consequently, to construct pre-asymptotic estimates of the bias terms, the unknowns, $\mu_{S-}^{(p+1)}$ and $\mu_{S+}^{(p+1)}$, are replaced by a q-th order ($p < q$) local polynomial estimate using a complementary bandwidth b. This leads to the pre-asymptotic feasible bias estimates $\hat{\mathcal{B}}_{\hat{\theta},p,q}(h,b) := \hat{\mathcal{B}}_{\hat{\theta}_+,p,q}(h,b) - \hat{\mathcal{B}}_{\hat{\theta}_-,p,q}(h,b)$ with

$$\hat{\mathcal{B}}_{\hat{\theta}_-,p,q}(h,b) := \mathbf{e}_0'\Gamma_{-,p}^{-1}(h)\vartheta_{-,p}(h)\frac{\mathbf{t}_{\hat{\theta}}'\dot{\mu}_{S-,q}^{(p+1)}(b)}{(p+1)!},$$
$$\hat{\mathcal{B}}_{\hat{\theta}_+,p,q}(h) := \mathbf{e}_0'\Gamma_{+,p}^{-1}(h,b)\vartheta_{+,p}(h)\frac{\mathbf{t}_{\hat{\theta}}'\dot{\mu}_{S+,q}^{(p+1)}(b)}{(p+1)!}$$

where $\dot{\mu}_{S-,q}^{(p+1)}(b)$ and $\dot{\mu}_{S+,q}^{(p+1)}(b)$ are the covariate-adjusted local polynomial estimates of the $(p+1)$-th derivatives using a polynomial of degree q.

As explained before, MSE-optimal bandwidths are too big to eliminate first-order bias. This motivates the need to recenter estimates using leading bias estimates for asymptotic estimations to remain valid. For a pair of bandwidth sequences $(\mathbf{h},\mathbf{b}) = (h_+,h_-,b_+,b_-)$, define the recentered statistic as

$$\hat{\theta}_{Centered,p,q}(\mathbf{h},\mathbf{b}) = \hat{\theta}(\mathbf{h}) - [h_+^{1+p}\hat{\mathcal{B}}_{\hat{\theta}_+,p,q}(h_+,b_+) - h_-^{1+p}\hat{\mathcal{B}}_{\hat{\theta}_-,p,q}(h_-,b_-)]$$

It follows from the formulas that

$$\hat{\theta}_{Centered,p,q}(\mathbf{h},\mathbf{b}) = \frac{1}{\sqrt{nh}}\left[\mathbf{t}_{\hat{\theta}}' \otimes \mathbf{e}_0'\left(\mathbf{P}_{+,,p,q}^{Centered}(h_+,b_+) - \mathbf{P}_{-,,p,q}^{Centered}(h_-,b_-)\right)\right]\mathbf{S},$$

where

$$\mathbf{P}_{-,p,q}^{Centered}(h,b) = \sqrt{h}\Gamma_{-,p}^{-1}(h)\left[\mathbf{R}_p(h)'\mathbf{K}_-(h) - (h/b)^{1+p}\vartheta_{-,p}(h)\mathbf{e}_{p+1}'\Gamma_{-,q}^{-1}(b)\mathbf{R}_q(b)'\mathbf{K}_-(b)\right]/\sqrt{n},$$
$$\mathbf{P}_{+,p,q}^{Centered}(h,b) = \sqrt{h}\Gamma_{+,p}^{-1}(h)\left[\mathbf{R}_p(h)'\mathbf{K}_+(h) - (h/b)^{1+p}\vartheta_{+,p}(h)\mathbf{e}_{p+1}'\Gamma_{+,q}^{-1}(b)\mathbf{R}_q(b)'\mathbf{K}_+(b)\right]/\sqrt{n}$$

Theorem A-2. *Suppose that regularity assumptions RA-1, RA-2, RA-3, RA-4 and RA-5 hold. Then, for any $\hat{\theta} \in \mathcal{F}_{p,m}$, consider the following standardized statistic*

$$\mathcal{T}_{S,\hat{\theta},q}(\mathbf{h},\mathbf{b}) = \frac{\hat{\theta}_{Centered,p,q}(\mathbf{h},\mathbf{b}) - \theta}{\sqrt{\mathbb{V}[\hat{\theta}_{Centered,p,q}(\mathbf{h},\mathbf{b})]}}$$

where

$$\mathbb{V}[\hat{\theta}_{Centered,p,q}(\mathbf{h},\mathbf{b})] = \frac{1}{nh_-}\mathcal{V}^{Centered}_{S-,\hat{\theta},p,q}(h_-,b_-) + \frac{1}{nh_+}\mathcal{V}^{Centered}_{S+,\hat{\theta},p,q}(h_+,b_+)$$

with

$$\mathcal{V}^{Centered}_{S-,\hat{\theta},p,q}(h_-,b_-) = (\mathbf{t}_{\hat{\theta}} \otimes \mathbf{e}_0)' \left[\mathbf{I}_{1+d} \otimes \mathbf{P}^{Centered}_{-,p,q}(h_-,b_-)\right] \Sigma_{\mathbf{S}} \left[\mathbf{I}_{1+d} \otimes \mathbf{P}^{Centered}_{-,p,q}(h_-,b_-)'\right] (\mathbf{t}_{\hat{\theta}} \otimes \mathbf{e}_0),$$
$$\mathcal{V}^{Centered}_{S+,\hat{\theta},p,q}(h_+,b_+) = (\mathbf{t}_{\hat{\theta}} \otimes \mathbf{e}_0)' \left[\mathbf{I}_{1+d} \otimes \mathbf{P}^{Centered}_{+,p,q}(h_+,b_+)\right] \Sigma_{\mathbf{S}} \left[\mathbf{I}_{1+d} \otimes \mathbf{P}^{Centered}_{+,p,q}(h_+,b_+)'\right] (\mathbf{t}_{\hat{\theta}} \otimes \mathbf{e}_0),$$

If $\rho \geq 1+q$, $p < q$, $n\min\{h_-,h_+\} \to \infty$, $nh_-^{2p+3}\max\{h_-^2, b_-^{2(q-p)}\} \to 0$, $nh_+^{2p+3}\max\{h_+^2, b_+^{2(q-p)}\} \to 0$ *and* $\limsup_{n\to\infty} \max\{h_-/b_-, h_+/b_+\} < \infty$, *then*

$$\mathcal{T}_{S,\hat{\theta},q} \xrightarrow{d} \mathcal{N}(0,1)$$

Proof. The proof follows as in Lemma SA-11, Calonico et al. (2019a), using Linderberg-Feller's triangular array central limit theorem (van der Vaart, 1998, p.20).

■

Theorem A-2 states that even when using MSE-optimal bandwidths, an asymptotically-normal pivot for $\hat{\theta}$ can be constructed by explicitly removing the first-order bias. Therefore, provided some consistent estimates $\hat{\Sigma}_{\mathbf{S}}$ of $\Sigma_{\mathbf{S}}$, it is possible to create an asymptotically normal pivotal statistic by replacing unknown quantities with their consistent estimators. These can be constructed using the traditional plug-in residuals estimator or a nearest-neighbor approach as developed by Calonico et al. (2014), which can be formulated to allow unrestricted forms of heteroskedasticity as well as clustered data.

Consequently, the distribution for the bias-corrected coefficient is:

Corollary A-2.1. *Under the assumptions of Theorem B.9, it follows that*

$$\mathcal{T}_{S,\hat{\tau}^{BC}_{PCRD},q}(\mathbf{h},\mathbf{b}) = \frac{\hat{\tau}^{BC}_{PCRD,Centered,p,q}(\mathbf{h},\mathbf{b}) - \tau_{PCRD}}{\sqrt{\hat{\mathbb{V}}[\hat{\tau}^{BC}_{PCRD,Centered,p,q}(\mathbf{h},\mathbf{b})]}} \xrightarrow{d} \mathcal{N}(0,1)$$

where

$$\hat{\tau}^{BC}_{PCRD,Centered,p,q}(\mathbf{h},\mathbf{b}) = \hat{\tau}^{BC}_{PCRD} - [h_+^{1+p}\hat{\mathcal{B}}_{\hat{\tau}^{BC}_{PCRD},+,p,q}(h_+,b_+) - h_-^{1+p}\hat{\mathcal{B}}_{\hat{\tau}^{BC}_{PCRD},-,p,q}(h_-,b_-)]$$

and

$$\hat{\mathbb{V}}[\hat{\tau}^{BC}_{PCRD,Centered,p,q}(\mathbf{h},\mathbf{b})] = \frac{1}{nh_-}\hat{\mathcal{V}}^{Centered}_{S-,\hat{\tau}^{BC}_{PCRD},p,q}(h_-,b_-) + \frac{1}{nh_+}\hat{\mathcal{V}}^{Centered}_{S+,\hat{\tau}^{BC}_{PCRD},p,q}(h_+,b_+)$$

with

$$\hat{\mathcal{V}}^{Centered}_{S-,\hat{\tau}^{BC}_{PCRD},p,q}(h_-,b_-) = (\hat{\mathbf{t}}_1\otimes\mathbf{e}_0)'\left[\mathbf{I}_{1+d}\otimes\mathbf{P}^{Centered}_{-,p,q}(h_-,b_-)\right]\hat{\Sigma}_{\mathbf{S}}\left[\mathbf{I}_{1+d}\otimes\mathbf{P}^{Centered}_{-,p,q}(h_-,b_-)'\right](\hat{\mathbf{t}}_1\otimes\mathbf{e}_0),$$
$$\hat{\mathcal{V}}^{Centered}_{S+,\hat{\tau}^{BC}_{PCRD},p,q}(h_+,b_+) = (\hat{\mathbf{t}}_1\otimes\mathbf{e}_0)'\left[\mathbf{I}_{1+d}\otimes\mathbf{P}^{Centered}_{+,p,q}(h_+,b_+)\right]\hat{\Sigma}_{\mathbf{S}}\left[\mathbf{I}_{1+d}\otimes\mathbf{P}^{Centered}_{+,p,q}(h_+,b_+)'\right](\hat{\mathbf{t}}_1\otimes\mathbf{e}_0),$$

B.10 Connection between the restricted and complete local models

Consider the simple scenario when there are two relevant variables we observe $\mathbf{Z}_d$, but not U_d. Furthermore, consider the modified covariate vectors:

$$\ddot{\mathbf{Z}}_d = \dot{\mathbf{Z}}_d - \bar{\mathbf{Z}} \in \mathbb{R}^{2m}, \quad \ddot{U}_d = \dot{U}_d - \bar{U} \in \mathbb{R}^2$$

where $\dot{U}_d$ is defined the same as $\dot{\mathbf{Z}}_d$ (See Appendix B.4). Then estimating the complete model (4.1) yields using bandwidths $\mathbf{h} = (h_+, h_-)$ and a local polynomial of degree p:

$$\begin{aligned}Y_d =& \hat{\ddot{\alpha}}_{Y,Full}(\mathbf{h}) + X_d\hat{\ddot{\tau}}_{Y,Full}(\mathbf{h}) + \hat{\ddot{\beta}}'_{Y-,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id}) + X_d\hat{\ddot{\beta}}'_{Y+,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id}) + \\ & X_d\ddot{\mathbf{Z}}'_d\hat{\ddot{\gamma}}_{+,Full}(\mathbf{h}) + (1-X_d)\ddot{\mathbf{Z}}'_d\hat{\ddot{\gamma}}_{-,Full}(\mathbf{h}) + X_d\ddot{U}_d\hat{\ddot{\rho}}_+(\mathbf{h}) + (1-X_d)\ddot{U}_d\hat{\ddot{\rho}}_-(\mathbf{h}) + \hat{\varepsilon}_{Full}(\mathbf{h})\end{aligned}$$

And the restricted model (4.2)

$$\begin{aligned}Y_{id} =& \hat{\ddot{\alpha}}_{Y,Res}(\mathbf{h}) + X_d\hat{\ddot{\tau}}_{Y,Res}(\mathbf{h}) + \hat{\ddot{\beta}}'_{Y-,Res}(\mathbf{h})\mathbf{r}_p(\Delta_{id}) + X_d\hat{\ddot{\beta}}'_{Y+,Res}(\mathbf{h})\mathbf{r}_p(\Delta_{id}) + \\ & X_d\ddot{\mathbf{Z}}'_d\hat{\ddot{\gamma}}_{+,Res}(\mathbf{h}) + (1-X_d)\ddot{\mathbf{Z}}'_d\hat{\ddot{\gamma}}_{-,Res}(\mathbf{h}) + \hat{\varepsilon}_{Res}(\mathbf{h})\end{aligned}$$

We are interested in studying the bias originating from unobserved confounding. For this purpose, define

$$\widehat{\text{Bias}}_p(\mathbf{h}) = \hat{\ddot{\tau}}_{Y,Full}(\mathbf{h}) - \hat{\ddot{\tau}}_{Y,Res}(\mathbf{h})$$

Proposition B.1. *In the situation described above,*

$$\widehat{Bias}_p(\mathbf{h}) = -\left[\hat{\ddot{\rho}}_{+,Full}(\mathbf{h})\hat{\ddot{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) - \hat{\ddot{\rho}}_{-,Full}(\mathbf{h})\hat{\ddot{\tau}}_{U-\bar{U},p,-}(\mathbf{h})\right]$$

Proof. The resticted model, implemented with bandwidths $\mathbf{h} = (h_-, h_+)$ and polynomial of degree p is obtained by solving:

$$\underset{b_-,b_+,\gamma}{argmin}\sum_{d=1}^{n}\left(Y_d - \mathbf{r}_{-,p}(\Delta_d)'\mathbf{b}_- - \mathbf{r}_{+,p}(\Delta_d)'\mathbf{b}_+ - \ddot{\mathbf{Z}}'_d\gamma\right)^2 K_{\mathbf{h}}(\Delta_d)$$

Furthermore, denote the oblique projection matrix into the column space of $\tilde{\mathbf{R}}_p(h)$ as

$$\mathbf{H}_{W,\tilde{\mathbf{R}}_p}(\mathbf{h}) = \tilde{\mathbf{R}}_p(\mathbf{h})(\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})\tilde{\mathbf{R}}_p(\mathbf{h}))^{-1}\tilde{\mathbf{R}}_p(\mathbf{h})'\mathbf{W}(\mathbf{h})$$

and consider the projection matrix into the complementary subspace as $\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h}) = (\mathbf{I}_n - \mathbf{H}_{W,\tilde{\mathbf{R}}_p}(\mathbf{h}))$. Using results derived in B.5, it follows that:

$$\hat{\hat{\gamma}}_{p,Res}(\mathbf{h}) = [\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}']^{-1}\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})Y$$

Now, by the idempotency property of $\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})$, remark that

$$\hat{\hat{\gamma}}_{p,Res}(\mathbf{h}) = [\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}']^{-1}\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})Y$$

Moreover, note that

$$\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h}) = [\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})]'\mathbf{W}(\mathbf{h})$$

So that

$$\hat{\hat{\gamma}}_{p,Res}(\mathbf{h}) = [[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]']^{-1}[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})Y]$$

Now, expressing the full model in matrix terms yields

$$Y = \tilde{\mathbf{R}}_p(\mathbf{h})\hat{\hat{\beta}}_{Y,p}(\mathbf{h}) + \ddot{\mathbf{Z}}\hat{\hat{\gamma}}_{p,Full}(\mathbf{h}) + \ddot{\mathbf{U}}\hat{\hat{\rho}}_{p,Full}(\mathbf{h}) + \hat{\varepsilon}_{Full}$$

Consequently,

$$\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})Y = \mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}\hat{\hat{\gamma}}_{p,Full}(\mathbf{h}) + \mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}\ddot{\mathbf{U}}\hat{\hat{\rho}}_{p,Full}(\mathbf{h})$$

and

$$\hat{\hat{\gamma}}_{p,Res}(\mathbf{h}) = \hat{\hat{\gamma}}_{p,Full}(\mathbf{h}) + [[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]']^{-1}[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{U}}]\hat{\hat{\rho}}_{p,Full}(\mathbf{h})$$

But observe that rearranging terms leads to

$$\begin{aligned}
&[[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]']^{-1}[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}]'\mathbf{W}(\mathbf{h})[\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{U}}] = \\
&[\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{Z}}']^{-1}\ddot{\mathbf{Z}}'\mathbf{W}(\mathbf{h})\mathbf{H}^c_{W,\tilde{\mathbf{R}}_p}(\mathbf{h})\ddot{\mathbf{U}} = \hat{\hat{\zeta}}_p(\mathbf{h})
\end{aligned}$$

where $\hat{\hat{\zeta}}_p(\mathbf{h})$ is the solution of the weighted least squares problem given by

$$\underset{b_-,b_+,\zeta}{argmin}\sum_{d=1}^{n}\left([U_d - \bar{U}] - \mathbf{r}_{-,p}(\Delta_d)'\mathbf{b}_- - \mathbf{r}_{+,p}(\Delta_d)'\mathbf{b}_+ - \ddot{\mathbf{Z}}'_d\zeta\right)^2 K_{\mathbf{h}}(\Delta_d)$$

Hence, we have demonstrated that

$$\hat{\hat{\gamma}}_{p,Res,+}(\mathbf{h}) = \hat{\hat{\gamma}}_{p,Full,+}(\mathbf{h}) + \hat{\hat{\rho}}_{p,Full,+}(\mathbf{h})\hat{\hat{\zeta}}_{p,+}(\mathbf{h})$$
$$\hat{\hat{\gamma}}_{p,Res,-}(\mathbf{h}) = \hat{\hat{\gamma}}_{p,Full,-}(\mathbf{h}) + \hat{\hat{\rho}}_{p,Full,-}(\mathbf{h})\hat{\hat{\zeta}}_{p,-}(\mathbf{h})$$

Now, following Theorem 5, and using the preceeding results

$$\begin{aligned}
\hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h}) &= \hat{\hat{\tau}}_{Y,p,+}(\mathbf{h}) - \frac{1}{2}(\hat{\hat{\tau}}_{\mathbf{Z},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z},p,-}(\mathbf{h}))'\hat{\hat{\gamma}}_{p,Res,+}(\mathbf{h}) \\
&= \hat{\hat{\tau}}_{Y,p,+}(\mathbf{h}) - \frac{1}{2}(\hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,-}(\mathbf{h}))'\hat{\hat{\gamma}}_{p,Full,+}(\mathbf{h}) \\
&- \frac{1}{2}(\hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,-}(\mathbf{h}))'[\hat{\hat{\rho}}_{p,Full,+}(\mathbf{h})\hat{\hat{\zeta}}_{p,+}(\mathbf{h})]
\end{aligned}$$

While

$$\begin{aligned}
\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) = \hat{\hat{\tau}}_{Y,p,+}(\mathbf{h}) &- \frac{1}{2}(\hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,-}(\mathbf{h}))'\hat{\hat{\gamma}}_{p,Full,+}(\mathbf{h}) \\
&- \frac{1}{2}(\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h}))\hat{\hat{\rho}}_{+,Full}
\end{aligned}$$

Then

$$\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h}) = \frac{\hat{\hat{\rho}}_{+,Full}(\mathbf{h})}{2}\left[(\hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z}-\bar{\mathbf{Z}},p,-}(\mathbf{h}))'\hat{\hat{\zeta}}_{p,+}(\mathbf{h}) - (\hat{\hat{\tau}}_{U-\bar{U},+}(\mathbf{h}) - \hat{\hat{\tau}}_{U-\bar{U},-}(\mathbf{h}))\right]$$

But using Theorem 5 again:

$$\frac{1}{2}\left[(\hat{\hat{\tau}}_{\mathbf{Z}-\bar{Z},p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{\mathbf{Z}-\bar{Z},p,-}(\mathbf{h}))'\hat{\hat{\zeta}}_{p,+}(\mathbf{h}) - (\hat{\hat{\tau}}_{U-\bar{U},+}(\mathbf{h}) - \hat{\hat{\tau}}_{U-\bar{U},-}(\mathbf{h}))\right] = -\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h})$$

Hence

$$\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h}) = -\hat{\hat{\rho}}_{+,Full}(\mathbf{h})\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h})$$

Likewise, by a similar procedure,

$$\hat{\hat{\tau}}_{Y,Full,p,-}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Res,p,-}(\mathbf{h}) = -\hat{\hat{\rho}}_{-,Full}(\mathbf{h})\hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h})$$

Hence, concluding the proof insofar as

$$\widehat{\mathrm{Bias}}_p(\mathbf{h}) = (\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h})) - (\hat{\hat{\tau}}_{Y,Full,p,-}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Res,p,-}(\mathbf{h}))$$

■

B.11 Charecterizing the omitted variable bias

Departing from Proposition B.1, consider the problem:

$$\underset{\tau_-,\tau_+,b_-,b_+,\zeta}{argmin} \sum_{d=1}^{n} \Big([U_d - \bar{U}] - \tau_-(1-X_d) - b_{-,1}((1-X_d)\times\Delta_d) - \cdots - b_{-,p}((1-X_d)\times\Delta_d)^p$$
$$- \tau_+ X_d - b_{+,1}(X_d\times\Delta_d) - \cdots - b_{+,p}(X_d\times\Delta_d)^p - \ddot{\mathbf{Z}}_d'\zeta \Big)^2 K_{\mathbf{h}}(\Delta_d)$$

The first-order conditions for τ_- and τ_+ associated to this problem yield

$$\sum_{d=1}^{n} \Big(X_d [U_d - \bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}} - \hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) X_d \Big) K(\Delta_d/h_+) = 0$$
$$\sum_{d=1}^{n} \Big((1-X_d) [U_d - \bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}} - \hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h})(1-X_d) \Big) K(\Delta_d/h_-) = 0$$

Where $K(\cdot)$ is the kernel function and $[U_d - \bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}$ is the residual of $(U_d - \bar{U})$ after controlling for the local functional form of Δ_p and removing the linear components assigned to $\ddot{\mathbf{Z}}$. More precisely,

$$[U_d - \bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}} = U_d - \bar{U} - \hat{\hat{\beta}}_{-,p,1}(\mathbf{h})((1-X_d)\times\Delta_p) - \cdots - \hat{\hat{\beta}}_{-,p,p}(\mathbf{h})((1-X_d)\times\Delta_p)^p$$
$$- \hat{\hat{\beta}}_{+,p,1}(\mathbf{h})(X_d\times\Delta_p) - \cdots - \hat{\hat{\beta}}_{+,p,p}(\mathbf{h})(X_d\times\Delta_p)^p - \ddot{\mathbf{Z}}_d'\hat{\hat{\zeta}}(\mathbf{h})$$

Then

$$\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) = \frac{\sum_{d=1}^{n} X_d Y_d^{\perp\Delta,\ddot{\mathbf{Z}}} K(\Delta_d/h_+)}{\sum_{d=1}^{n} X_d K(\Delta_d/h_+)}; \qquad \hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h}) = \frac{\sum_{d=1}^{n} (1-X_d) Y_d^{\perp\Delta,\ddot{\mathbf{Z}}} K(\Delta_d/h_-)}{\sum_{d=1}^{n} (1-X_d) K(\Delta_d/h_-)}$$

Hence note that

$$\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) = \frac{\widehat{\mathrm{Cov}}_K(X_d[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}, X_d)[h_+]}{\widehat{\mathrm{Var}}_K(X_d)[h_+]}$$

Where

$$\widehat{\mathrm{Cov}}_K(A_d, B_d)[h] = \sum_{d=1}^{n} (A_d - \hat{\mathbb{E}}_K(A_d)[h])(B_d - \hat{\mathbb{E}}_K(B_d)[h])' K(\Delta_d/h)$$

$\widehat{\mathrm{Var}}_K(A_d)[h] = \widehat{\mathrm{Cov}}_K(A_d, A_d)[h]$, and

$$\hat{\mathbb{E}}_K(A_d)[h] = \left(\sum_{d=1}^{n} A_d K(\Delta_d/h) \right) \Big/ \sum_{d=1}^{n} K(\Delta_d/h)$$

Similarly

$$\hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h}) = \frac{\widehat{\text{Cov}}_K((1-X_d)[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}, 1-X_d)[h_-]}{\widehat{\text{Var}}_K(1-X_d)[h_-]}$$

In a similar fashion, to estimate $\ddot{\rho}$ we consider the problem:

$$\underset{b_-,b_+,\gamma,\rho}{argmin} \sum_{d=1}^{n} \left(Y_d - (1-X_d)\mathbf{r}_p(\Delta_d)'\mathbf{b}_- - X_d\mathbf{r}_p(\Delta_d)'\mathbf{b}_+ - \ddot{\mathbf{Z}}'\gamma - \ddot{\mathbf{U}}'\rho\right)^2 K_{\mathbf{h}}(\Delta_d)$$

Where, by the Frisch-Waugh-Lovell Theorem, it follows that

$$\hat{\hat{\rho}}_{+,Full}(\mathbf{h}) = \frac{\widehat{\text{Cov}}_K(X_d Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}}, X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}$$
$$\hat{\hat{\rho}}_{-,Full}(\mathbf{h}) = \frac{\widehat{\text{Cov}}_K((1-X_d)Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}}, (1-X_d)[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_-]}{\widehat{\text{Var}}_K((1-X_d)[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_-]}$$

where

$$Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}} = Y_d - (1-X_d)r_p(\Delta_d)'\hat{\hat{\beta}}_{Y,-}(\mathbf{h}) - X_d r_p(\Delta_d)'\hat{\hat{\beta}}_{Y,+}(\mathbf{h}) - \ddot{\mathbf{Z}}_d'\hat{\hat{\gamma}}_{p,Full}(\mathbf{h})$$

and

$$[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}} = [U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}} - \hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h})X_d - \hat{\hat{\tau}}_{U-\bar{U},p,-}(\mathbf{h})(1-X_d)$$

Now

$$\begin{aligned}
\hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) &= \hat{\hat{\rho}}_{+,Full}(\mathbf{h})\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h}) \\
&= \left(\frac{\widehat{\text{Cov}}_K(X_d Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}}, X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}\right)\left(\frac{\widehat{\text{Cov}}_K(X_d[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}, X_d)[h_+]}{\widehat{\text{Var}}_K(X_d)[h_+]}\right) \\
&= \left(\frac{\widehat{\text{Corr}}_K(X_d Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}}, X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]\sqrt{\widehat{\text{Var}}_K(X_d Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}}{\sqrt{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}}\right) \\
&\times \left(\frac{\widehat{\text{Corr}}_K(X_d[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}, X_d)[h_+]\sqrt{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}})[h_+]}}{\sqrt{\widehat{\text{Var}}_K(X_d)[h_+]}}\right)
\end{aligned}$$

Now observe that

$$\frac{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}})[h_+]} = 1-\hat{\hat{\tau}}_{U-\bar{U},p,+}(\mathbf{h})\frac{\widehat{\text{Var}}_K(X_d)[h_+]}{\widehat{\text{Var}}_K(X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}})[h_+]}$$
$$= 1-\left(\widehat{\text{Corr}}_K(X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}},X_d)[h_+]\right)^2$$

Therefore

$$\hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h})-\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) = \sqrt{\frac{\widehat{\text{Var}}_K(X_dY_d^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_+]}{\widehat{\text{Var}}_K(X_d)[h_+]}}\left(\frac{R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,+}[h_+]R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h_+]}{\sqrt{1-R^2_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h_+]}}\right)$$

where

$$R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,+}[h] = \widehat{\text{Corr}}_K(X_dY_d^{\perp X,\Delta,\ddot{\mathbf{Z}}},X_d[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h]$$
$$R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h] = \widehat{\text{Corr}}_K(X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}},X_d)[h]$$

Ultimately, previous results mean that the bias caused by leaving out a potential confounder depends on two unknown quantities: $R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,+}[h]$ and $R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,+}[h]$. To simplify notation, let

$$\hat{\hat{\tau}}_{Y,Res,p,+}(\mathbf{h})-\hat{\hat{\tau}}_{Y,Full,p,+}(\mathbf{h}) = \hat{\mathcal{C}}_+[h_+]\text{BF}_+[h_+]$$

where

$$\text{BF}_+(h) = \left(\frac{R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,+}[h]R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h]}{\sqrt{1-R^2_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h]}}\right)$$

and

$$\hat{\mathcal{C}}_+[h] = \sqrt{\frac{\widehat{\text{Var}}_K(X_dY_d^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h]}{\widehat{\text{Var}}_K(X_d)[h]}}$$

Further note that correlations are symmetric, which means that

$$R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,+}[h] = \widehat{\text{Corr}}_K(X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}},X_d)[h] = \widehat{\text{Corr}}_K(X_d,X_d[U_d-\bar{U}]^{\perp \Delta,\ddot{\mathbf{Z}}})[h] = R_{X\sim U-\bar{U}|\mathbf{Z},\Delta,+}[h]$$

Mimicking this analysis in the left-side yields

$$\hat{\hat{\tau}}_{Y,Res,p,-}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Full,p,-}(\mathbf{h}) = \sqrt{\frac{\widehat{\text{Var}}_K((1-X_d)Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h_-]}{\widehat{\text{Var}}_K(1-X_d)[h_-]}} \left(\frac{R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,-}[h_-] R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,-}[h_-]}{\sqrt{1-R^2_{U-\bar{U}\sim X|\mathbf{Z},\Delta,-}[h_-]}} \right)$$

where

$$R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,-}[h] = \widehat{\text{Corr}}_K((1-X_d)Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}}, (1-X_d)[U_d-\bar{U}]^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h]$$
$$R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,-}[h] = -\widehat{\text{Corr}}_K((1-X_d)[U_d-\bar{U}]^{\perp\Delta,\ddot{\mathbf{Z}}}, X_d)[h]$$

Analogously, by symmetry,

$$\hat{\hat{\tau}}_{Y,Res,p,-}(\mathbf{h}) - \hat{\hat{\tau}}_{Y,Full,p,-}(\mathbf{h}) = \hat{\mathcal{C}}_-[h_-]\text{BF}_-(h_-)$$

where

$$\text{BF}_-(h) = \left(\frac{R_{Y\sim U-\bar{U}|\mathbf{Z},\Delta,X,-}[h] R_{U-\bar{U}\sim X|\mathbf{Z},\Delta,-}[h]}{\sqrt{1-R^2_{U-\bar{U}\sim X|\mathbf{Z},\Delta,-}[h]}} \right)$$

and

$$\hat{\mathcal{C}}_-[h] = \sqrt{\frac{\widehat{\text{Var}}_K((1-X_d)Y_d^{\perp X,\Delta,\ddot{\mathbf{Z}}})[h]}{\widehat{\text{Var}}_K(1-X_d)[h]}}$$

As a result, the result can be succinctly stated as

$$\widehat{\text{Bias}}(\mathbf{h}) = -\left(\hat{\mathcal{C}}_+[h_+]\text{BF}_+(h_+) + \hat{\mathcal{C}}_-[h_-]\text{BF}_-(h_-)\right)$$

B.12 Bounding estimated effect using observed covariates

Let Z_{idl} be the benchmark characteristic and denote by $\mathbf{Z}_{id}^{-l}$ the observed trait vector excluding attribute Z_{idl}. Now let

$$\kappa_X^+ := \frac{R^2_{X\sim U-\bar{U}|\mathbf{Z}^{-l},\Delta,+}[h_+]}{R^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}; \quad \kappa_Y^+ := \frac{R^2_{Y\sim U-\bar{U}|\mathbf{Z}^{-l},\Delta,X,+}[h_+]}{R^2_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]}$$

Following Appendix B.2 of Cinelli and Hazlett (2020b), it is possible to use the recursive nature of correlations to bound the $|\text{BF}^+|$ using κ_X^+, κ_Y^+ and some estimated quantities in the data. In particular,

$$|R_{X\sim U-\bar{U}|\mathbf{Z},\Delta,+}[h_+]| = \sqrt{\kappa_X^+}|\Lambda_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]|$$
$$|R_{X\sim U-\bar{U}|\mathbf{Z},\Delta,+}[h_+]| \leq \Xi(\kappa_X^+,\kappa_Y^+)|\Lambda_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]|$$

where

- $$\Lambda_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+] = \frac{R_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}{\sqrt{1-R^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}}$$
- $$\Lambda_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+] = \frac{R_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]}{\sqrt{1-R^2_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]}}$$
- $$\Lambda_{\kappa_X^+}[h_+] = \frac{\sqrt{\kappa_X^+}R_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}{\sqrt{1-\kappa_X^+R^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}}$$
- $$\Xi^+(\kappa_X^+,\kappa_Y^+) = \frac{\sqrt{\kappa_Y^+}+|\Lambda_{\kappa_X^+}[h_+]\times\Lambda_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]|}{\sqrt{1-\Lambda^2_{\kappa_X^+}[h_+]\times\Lambda^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}}$$

are all known quantities. Consequently,

$$|\mathrm{BF}^+|\leq \Xi^+(\kappa_X^+,\kappa_Y^+)\left|\Lambda_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]\right|\left(\frac{\sqrt{\kappa_X^+}|\Lambda_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]|}{\sqrt{1-\kappa_X^+\Lambda^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]}}\right) = \overline{\mathrm{BF}}^+(\kappa_X^+,\kappa_Y^+)[h_+]$$

and this bound is tight, which means it is the smallest possible bound for this quantity.

Moreover since $R^2_{X\sim U-\bar{U}|\mathbf{Z},\Delta,+}[h_+]\leq 1$, it follows that

$$0\leq\kappa_X^+\leq 1/\Lambda^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]|$$

Likewise, by considering

$$\kappa_X^- := \frac{R^2_{X\sim U-\bar{U}|\mathbf{Z}^{-l},\Delta,-}[h_-]}{R^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,-}[h_-]};\quad \kappa_Y^- := \frac{R^2_{Y\sim U-\bar{U}|\mathbf{Z}^{-l},\Delta,X,-}[h_-]}{R^2_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,-}[h_-]}$$

It is possible to bound $|\mathrm{BF}_-|$ as follows:

$$|\mathrm{BF}^-|\leq \Xi(\kappa_X^-,\kappa_Y^-)\Big|\Lambda_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,-}[h_-]\Big|\left(\frac{\sqrt{\kappa_X^-}|\Lambda_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,-}[h_-]|}{\sqrt{1-\kappa_X^-\Lambda^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,-}[h_-]}}\right)=\overline{\mathrm{BF}}^-(\kappa_X^-,\kappa_Y^-)[h_-]$$

where the quantities indexed with $-$ are defined the same as their conterparts with $+$. Similarly, we know that

$$0\leq \kappa_X^-\leq 1/\Lambda^2_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,-}[h_-]$$

Now, from the triangular inequality

$$|\widehat{\mathrm{Bias}}(\mathbf{h})|\leq \hat{\mathcal{C}}_+[h_+]|\mathrm{BF}_+(h_+)|+\hat{\mathcal{C}}_-[h_-]|\mathrm{BF}_-(h_-)|$$

where equality is achieved in the likely scenario where $\mathrm{BF}_+(h_+)$ and $\mathrm{BF}_-(h_-)$ have the same sign[22]. In conclusion

$$|\widehat{\mathrm{Bias}}(\mathbf{h})|\leq \hat{\mathcal{C}}_+[h_+]\times\overline{\mathrm{BF}}^+(\kappa_X^+,\kappa_Y^+)[h_+]+\hat{\mathcal{C}}_-[h_-]\times\overline{\mathrm{BF}}^-(\kappa_X^-,\kappa_Y^-)[h_-]=\overline{\mathrm{Bias}}(\kappa_X^+,\kappa_Y^+,\kappa_X^-,\kappa_Y^-)$$

B.13 Extensions to the sensitivity analysis.

B.13.1 Multiple omitted confounders

Section 4.1 assumes that there only exists a single omitted confounder. I will show that this simplified framework covers also setting with multiple confounders. Let $\mathbf{U}_d$ be a *vector* of unobserved multiple confounders. In this scenario, the complete model would be

$$\begin{aligned}Y_d=&\hat{\hat{\alpha}}_{Y,Full}(\mathbf{h})+X_d\hat{\hat{\tau}}_{Y,Full}(\mathbf{h})+\hat{\hat{\beta}}'_{Y-,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id})+X_d\hat{\hat{\beta}}'_{Y+,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id})+\\&X_d\ddot{\mathbf{Z}}'_d\hat{\hat{\gamma}}_{+,Full}(\mathbf{h})+(1-X_d)\ddot{\mathbf{Z}}'_d\hat{\hat{\gamma}}_{-,Full}(\mathbf{h})+X_d\ddot{\mathbf{U}}'_d\hat{\hat{\rho}}_+(\mathbf{h})+(1-X_d)\ddot{\mathbf{U}}'_d\hat{\hat{\rho}}_-(\mathbf{h})+\hat{\varepsilon}_{Full}(\mathbf{h})\end{aligned}$$

Now consider the single variable $U_d^*=(X_d\hat{\hat{\rho}}_+(\mathbf{h})+(1-X_d)\hat{\hat{\rho}}_-)'\mathbf{U}_d$. Then the model

$$\begin{aligned}Y_d=&\hat{\hat{\alpha}}_{Y,Full}(\mathbf{h})+X_d\hat{\hat{\tau}}_{Y,Full}(\mathbf{h})+\hat{\hat{\beta}}'_{Y-,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id})+X_d\hat{\hat{\beta}}'_{Y+,Full}(\mathbf{h})\mathbf{r}_p(\Delta_{id})+\\&X_d\ddot{\mathbf{Z}}'_d\hat{\hat{\gamma}}_{+,Full}(\mathbf{h})+(1-X_d)\ddot{\mathbf{Z}}'_d\hat{\hat{\gamma}}_{-,Full}(\mathbf{h})+X_d(U_d^*-\bar{U}^*)+(1-X_d)(U_d^*-\bar{U}^*)\end{aligned}$$

yields the same $\hat{\hat{\tau}}_{Y,Full}(\mathbf{h})$. However, as Cinelli and Hazlett (2020b) notes,

[22]This would be the case if the unobserved trait correlates differently with the outcome depending on the politician's type.

$$|R_{\mathbf{U}-\bar{\mathbf{U}}\sim X|\mathbf{Z},\Delta,+}[h]| \leq |R_{U^*-\bar{U}^*\sim X|\mathbf{Z},\Delta,+}[h]|$$
$$|R_{\mathbf{U}-\bar{\mathbf{U}}\sim X|\mathbf{Z},\Delta,-}[h]| \leq |R_{U^*-\bar{U}^*\sim X|\mathbf{Z},\Delta,-}[h]|$$

so that the bias that is generated by omitting $\mathbf{U}_d$ must be less than equal than omitting U_d^*. Thus, by assuming the existence of a single unobserved feature U_d^*, we are still providing valid upper bounds in the presence of more than one omitted confounder.

B.13.2 Benchmarking with a group of covariates

To generalize the bounds to multiple covariates, consider $\mathbf{Z}_{(l_1,\cdots,l_r)} = \{Z_{l_1},\cdots,Z_{l_r}\}$, where $1 \leq l_1 < l_2 < \cdots < l_r \leq m$ be a subset of $\mathbf{Z}$ of size r. This set can be as big as to contain all observed variables (in which case $r = m$), or as small as to contain a single variable ($r = 1$). Denote by $\mathbf{Z}_{-(l_1,\cdots,l_r)}$ the set containing all the remaining traits not contained in $\mathbf{Z}_{(l_1,\cdots,l_r)}$. Finally, denote by $\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)} = \{Z_{l_1} - \bar{Z}_{l_1},\cdots,Z_{l_r} - \bar{Z}_{l_r}\}$.

Let

$$\kappa_X^+ := \frac{R^2_{X\sim U-\bar{U}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,+}[h_+]}{R^2_{X\sim\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,+}[h_+]}; \quad \kappa_Y^+ := \frac{R^2_{Y\sim U-\bar{U}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,X,+}[h_+]}{R^2_{Y\sim\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,X,+}[h_+]}$$

As in the original case, we compare how much more of the variance of the treatment or the outcome variable does the omitted confounder explain relative to the whole set of covariates (after controlling for the left-out traits). This makes it possible to produce the same bounds as in the single covariate case, simply replacing $R_{X\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,+}[h_+]$ with $R_{X\sim\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,+}[h_+]$, and $R_{Y\sim Z_{idl}-\bar{Z}_l|\mathbf{Z}^{-l},\Delta,X,+}[h_+]$ with $R_{Y\sim\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,X,+}[h_+]$.

Furthermore, one can use the recursive definition of partial correlations to calculate these quantities. For instance,

$$\begin{aligned}
1 - R^2_{X\sim\ddot{\mathbf{Z}}_{(l_1,\cdots,l_r)}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,+}[h_+] = &(1 - R^2_{X\sim Z_{l_1}-\bar{Z}_{l_1}|\mathbf{Z}_{-(l_1)},\Delta,+}[h_+]) \\
&\times (1 - R^2_{X\sim Z_{l_2}-\bar{Z}_{l_2}|\mathbf{Z}_{-(l_1,l_2)},\Delta,+}[h_+]) \\
&\times (1 - R^2_{X\sim Z_{l_3}-\bar{Z}_{l_3}|\mathbf{Z}_{-(l_1,l_2,l_3)},\Delta,+}[h_+]) \\
&\times \cdots \\
&\times (1 - R^2_{X\sim Z_{l_r}-\bar{Z}_{l_r}|\mathbf{Z}_{-(l_1,\cdots,l_r)},\Delta,+}[h_+])
\end{aligned}$$

The same applies to the quantities indexed by " - ". The formal proof of this equivalence is available in Cinelli and Hazlett (2020b, p.8).

C Appendix - Extensions and additional comments on the sensitivity analysis.

In this Appendix, I discuss additional procedures and perspectives to the sensitivity analysis described in Section 4.

C.1 Validity of the bounding exercise

In general terms, the bounding procedure enables to assert for any given tuple $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$ that

$$\hat{\hat{\tau}}_{Y,Res} - \overline{\text{Bias}}(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-) \leq \hat{\hat{\tau}}_{Y,Full} \approx \tau_{PCRD} \leq \hat{\hat{\tau}}_{Y,Res} + \overline{\text{Bias}}(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$$

Although a general result, the multiplicity of dimensions makes interpretation difficult. Because of this, Cinelli and Hazlett (2020a), advocate reducing to a single dimension by assuming that $\kappa_X^+ = \kappa_Y^+ = \kappa_X^- = \kappa_Y^- = \kappa$.

While this equality is likely false in any empiric application, it nevertheless gives valid bounds to the bias. However, to achieve this, we must make precise what being "κ times as strong" as the benchmark means. Rigorously, we can define κ as the strongest relative strength in any of the dimensions studied. More precisely, let $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$ be the true relative strengths define

$$\kappa = \max\{\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-\}$$

Since $\overline{\text{Bias}}(\cdot)$ is increasing on all of its arguments, we have

$$\overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa) \leq \overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa)$$

so that

$$\hat{\hat{\tau}}_{Y,Res} - \overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa) \leq \hat{\hat{\tau}}_{Y,Full} \approx \tau_{PCRD} \leq \hat{\hat{\tau}}_{Y,Res} + \overline{\text{Bias}}(\kappa, \kappa, \kappa, \kappa)$$

While valid, these bounds may be excessively large, making sensitivity bounds overly conservative. However, without additional information about the vector $(\kappa_X^+, \kappa_Y^+, \kappa_X^-, \kappa_Y^-)$, this approach furnishes valid bounds irrespective of the interrelation of these variables.

C.2 Assessing relative strength with post-treatment bias.

The key to the bounding procedure presented in Section 4.2 is to benchmark an omitted confounder's strength via its relative strength with respect to some observed pre-determined variables (consider a single benchmark Z_{id} for simplicity). More precisely, a researcher could be able to make sense of the quantities $\kappa_X^\pm$ and $\kappa_Y^\pm$ defined as

$$\kappa_X^+ := \left(\frac{R^+_{X_d\sim U_d-\bar{U}}}{R^+_{X_d\sim Z_d-\bar{Z}}}\right)^2; \quad \kappa_Y^+ := \left(\frac{R^+_{Y_d\sim U_d-\bar{U}|X_d}}{R^+_{Y_d\sim Z_d-\bar{Z}|X_d}}\right)^2$$

These quantities have a clear interpretation. For instance, κ_X^+ represents how much more of the variance of X_d near the threshold does $U_d-\bar{U}$ explain relative to the benchmark difference $Z_d-\bar{Z}$. Alternatively,

$$\kappa_X^+ \approx \frac{\text{Corr}^2(X_d, X_dU_d|\Delta_d=0)}{\text{Corr}^2(X_d, X_dZ_d|\Delta_d=0)}$$

Similarly, κ_Y^+ expresses a similar ratio but accounting for how much $U_d-\bar{U}$ explains the outcome at the right of the threshold, relative to $Z_d-\bar{Z}$. Therefore, it can be expressed as

$$\kappa_Y^+ \approx \frac{\text{Corr}^2(X_dY_d, X_dU_d|\Delta_d=0)}{\text{Corr}^2(X_dY_d, X_dZ_d|\Delta_d=0)}$$

However, notice that conditioning by a close race can distort these ratios relative to the population ones given by:

$$\tilde{\kappa}_X^+ \approx \frac{\text{Corr}^2(X_d, X_dU_d)}{\text{Corr}^2(X_d, X_dZ_d)}; \qquad \tilde{\kappa}_Y^+ \approx \frac{\text{Corr}^2(X_dY_d, X_dU_d)}{\text{Corr}^2(X_dY_d, X_dZ_d)}$$

The problem with this is that while expert knowledge can provide a good sense of $\tilde{\kappa}_X^\pm$ and $\tilde{\kappa}_Y^\pm$, the bias induced by conditioning can make the quantities we need to reason about, $\kappa_X^\pm$ and $\kappa_Y^\pm$, very different. To illustrate this, I will show how conditioning on a close race affects the ratio $\kappa_X^+/\tilde{\kappa}_X^+$. The intuition carries over to κ_X^- and $\kappa_Y^\pm$.

Suppose U_d is relevant for determining both Y_d and Δ_d, but Z_d is only relevant for Y_d. In this case, conditioning on $\Delta_d=0$ does not induce compensating differentials on Z_d. Therefore,

$$\text{Corr}^2(X_dY_d, X_dZ_d|\Delta_d=0) = \text{Corr}^2(X_dY_d, X_dZ_d)$$

and

$$\frac{\kappa_X^+}{\tilde{\kappa}_X^+} = \left(\frac{\text{Corr}(X_d, X_dU_d|\Delta_d=0)}{\text{Corr}(X_d, X_dU_d)}\right)^2$$

Thus, the question of whether $\tilde{\kappa}_X^+$ is greater or lower than our reference point κ_X^+ in this example can be reduced to evaluating whether conditioning in a close race accentuates the correlation of U_d with X_d. This occurs when differences in unobservables are accentuated by focusing on close elections. Since this is most likely the case, this example would suggest that we should always expect $\tilde{\kappa}_X^+ > \kappa_X^+$, meaning estimates will be even more sensible than we think.

Nonetheless, this pattern does not always hold in a more realistic scenario. For instance, if we allow Z_d to influence Δ_d. In this case,

$$\frac{\kappa_X^+}{\tilde{\kappa}_X^+} = \left(\frac{\text{Corr}(X_d, X_d U_d | \Delta_d = 0)}{\text{Corr}(X_d, X_d U_d)} \frac{\text{Corr}(X_d, X_d Z_d)}{\text{Corr}(X_d, X_d Z_d | \Delta_d = 0)} \right)^2$$

Now, the question is whether the benchmark attribute or the unobservable was more distorted by conditioning on a close election. In other words, one must determine whether

$$\frac{\text{Corr}(X_d, X_d U_d | \Delta_d = 0)}{\text{Corr}(X_d, X_d U_d)} \overset{?}{>} \frac{\text{Corr}(X_d, X_d Z_d | \Delta_d = 0)}{\text{Corr}(X_d, X_d Z_d)}$$

In other words, if the unobservables were more distorted by conditioning than the benchmark characteristic, then $\kappa_X^+ > \tilde{\kappa}_X^+$. Conversely, if the benchmark characteristic was more affected than the unobservables, then $\tilde{\kappa}_X^+ > \kappa_X^+$. Moreover, it could even be the case where $\tilde{\kappa}_X^+ = \kappa_X^+$ despite the existence of compensating differentials.

While not much can be done about how conditioning affects the unobserved variable, a researcher still has the power to choose an appropriate benchmark. Specifically, By selecting a benchmark that has been largely affected by conditioning, the ratio $\kappa_X^+ / \tilde{\kappa}_X^+$ is forced to be smaller. This brings the global benchmarks closer to the conditional ones, even when distortions in the unobservable are potentially as large.