Identification of Treatment Effects in Regression Discontinuity Designs with Measurement Error*

Ping Yu
University of Auckland

First Version: February 2011
This Version: June 2013

Abstract

Precise measurement of the forcing variable in regression discontinuity designs is critical to precise evaluation of treatment effects. Such evaluation can be sensitive to measurement errors, which are prevalent in many applications of regression discontinuity designs. The present paper studies identification of treatment effects using local polynomial estimation in the presence of measurement error. The main findings are as follows. In sharp designs, when the measurement error is fixed, the treatment effect can be identified in some special cases if the treatment is based on the contaminated forcing variable, but cannot be identified if the treatment is based on the error free forcing variable. If the measurement error is local to zero, the treatment effect can be identified with a small extra bias and without efficiency loss if the treatment is based on the contaminated forcing variable; the treatment effect can be identified with efficiency loss and a large bias if the treatment is based on the error free forcing variable and the treatment status can be observed; the treatment effect cannot be identified if the treatment is based on the error free forcing variable and the treatment status cannot be observed unless the measurement error is extremely small. The results are extended to fuzzy designs. Monte Carlo results confirm the theoretical analysis.

KEYWORDS: Regression Discontinuity Design, Measurement Error, Identifiability, Local Polynomial Estimator, Essential Heterogeneity
JEL-CLASSIFICATION: C13, C14, C21

*Email: p.yu@auckland.ac.nz. I would like to thank Stephen Karlson, Tommaso Proietti, Andrew Whitby and seminar participants at University of Auckland, Northern Illinois University and NZESG for helpful comments. Special thanks are given to Peter Phillips for carefully modifying part of the paper and the Reserve Bank of New Zealand (RBNZ) for sponsoring the RBNZ-NZESG Award to this paper.

1 Introduction

Regression discontinuity designs (RDDs) are quasi-experimental designs where treatment is determined by whether an observed *forcing* (*running* or *assignment* or *selection*) variable crosses a known threshold. As shown in Hahn et al. (2001), in the left and right neighborhoods of the threshold, the treatment is assigned as if in a randomized experimental design. So the individuals marginally below the threshold represent a valid counterfactual for the treated group just above the threshold. Given that a genuinely randomized design is rare in social science, RDDs have received much attention since their introduction by Thistlethwaite and Campbell (1960). See Cook (2008) for a history of RDDs in three academic disciplines; see Imbens and Lemieux (2008), van der Klaauw (2008) and Lee and Lemieux (2010) for excellent reviews on up-to-date theoretical developments and applications.

As emphasized by Lee and Lemieux (2010), the key element in RDDs is that treatment is determined solely by the forcing variable. In consequence, precise measurement of the forcing variable is critical to all empirical applications of RDD methodology. As stated in their Section 7.1,

> "even if there is perfect compliance of the discontinuous rule, it may be that the researcher does not directly observe the assignment variable, but instead possesses a slightly noisy measure of the variable. Understanding the effects of this kind of measurement error could further expand the applicability of RDDs".

A prototypical example of RDDs with measurement error is the pretest-posttest design in compensatory educational programs considered in Goldberger (2008); see also Matsudaira (2008). In that example, pupils are selected to take a compensatory program based on whether the forcing variable crosses the threshold. So the treatment is the program, and the outcome is the posttest result. The forcing variable can take two forms. In the first form, it is the pretest result. So the forcing variable can be observed by both the program designer and the econometrician. In the second form, it is true ability, so the pretest result is only a contaminated measure of the forcing variable, and can be precisely observed only by the program designer; the econometrician can only observe a noisy measure. In either case, we are interested in the treatment effect on pupils with *true* ability just above the threshold relative to those whose ability is just below the threshold. So the usual estimates must contain some bias for the true treatment effect since some information is unavailable to the econometrician. Even in the classical application of RDDs in U.S. house elections by Lee (2008), there can be measurement error. In that example the forcing variable is the democratic vote share margin of victory, the treatment is to be the incumbent party, and the outcome is the next election result. The closer the two parties' contesting powers, the more likely there will be a measurement error on the true ballot counting process – consider, for instance, the presidential election between Kennedy/Nixon or Bush/Gore. But as argued in Lee (2008), the data in such elections are critical to evaluate some kind of average treatment effect, and of course are critical in evaluating the true treatment effect - the treatment effect based on the *true* ballot counting process. More examples can be found in Pei (2011) where he emphasizes that measurement errors may most likely arise when the forcing variable is based on survey data; see Bound et al. (2001) for a summary on measurement error in survery data.

There are some preliminary results on the estimation of treatment effects in this environment. To facilitate exposition, we define some notations. Following Trochim (1984), when there is no measurement error, we classify RDDs into sharp designs and fuzzy designs. In both designs, the outcome equation is

$$y = m_\pi(x^*) + \alpha_\pi(x^*)\,D + \varepsilon \equiv m(x^*) + \varepsilon,$$

where the forcing variable x^* is some basic determinant of the outcome, D is the treatment status, $\varepsilon = \varepsilon_d$

when $D = d$, and $E[\varepsilon_d|x^*] = 0$, for $d = 0, 1$. The treatment effect is $\alpha_\pi(x^*) + \varepsilon_1 - \varepsilon_0$. The difference between the sharp design and the fuzzy design is whether D is a deterministic function of x^* or not. If $D = d^*_\pi \equiv 1(x^* \geq \pi)$, then we get a sharp design; if $D = p_\pi(x^*) + \beta_\pi(x^*)d^*_\pi + \eta$, where $E[\eta|x^*] = 0$, $p_\pi(x^*) \neq 0$ and $\beta_\pi(\pi) \neq 0$, then we have a fuzzy design. In both cases, we are interested in the average treatment effect at $x^* = \pi$, that is, $\alpha_\pi(\pi)$. If x^* is contaminated, we observe only $x = x^* + u$ instead of x^*, where u is a classical measurement error independent of all other random variables.[1] When there is a measurement error in x^*, the design is termed the *nonequivalent group design* (NEGD) in Cook and Campbell (1979). For future reference, we use the following classification. "Case 1" refers to the sharp design where the forcing variable is x (that is, $D = d_\pi \equiv 1(x \geq \pi)$), "Case 2" is the sharp design where the forcing variable is x^* with the treatment status d^*_π observed, "Case 3" is the sharp design where the forcing variable is x^* with d^*_π unobserved, "Case 4" is the fuzzy design where the forcing variable is x, "Case 5" is the fuzzy design where the forcing variable is x^*, and the "Oracle Case" is the design without measurement error. A closely related concern with Case 1 in the literature is the so-called *heaping* problem in x^*; see, e.g., Almond et al. (2010), Almond et al. (2011), Barreca et al. (2011a) and Barreca et al. (2011b). In a heaping problem, x^* is recorded in multiples of some basic unit. Our analysis provides a benchmark for understanding estimation bias in random heaping.[2] Examples of Case 2 include de la Mata (2011), Hullegie and Klein (2010), Koch (2010) and Schanzenbach (2009). The first form of the example in Goldberger (2008) and Lee (2008) both belong to Case 1, and the second form of Golberger (2008) falls in Case 2. Case 3 is logically possible but rarely happens in practice. We study it only because it is the building block of Case 5.

When there is a measurement error in x^*, all the existing literature considers the sharp design with $m_\pi(\cdot)$ and $\alpha_\pi(\cdot)$ being linear and the distributions of x^*, ε and u being known (especially as Gaussian). The main results in these special cases are summarized as follows. First, in Case 1, if $\alpha_\pi(\cdot)$ is a constant, then the least squares estimator (LSE) of $\alpha_\pi(\pi)$ obtained by regressing y on x and d_π is unbiased although the intercept and slope in $m_\pi(\cdot)$ are biased. In other words, the bias introduced by measurement error is completely absorbed in $m_\pi(\cdot)$ and the unbiasedness of $\alpha_\pi(\pi)$ remains as long as the treatment assignment is perfectly controlled; see Cappelleri et al. (1991) and Goldberger (2008). Second, in Case 1, if the slope of $\alpha_\pi(\cdot)$ is not zero, then the LSE of $\alpha_\pi(\pi)$ is biased but converges to the treatment effect at another point on the support of x^*; see Trochim et al. (1991). Third, in Case 2, the LSE of $\alpha_\pi(\pi)$ by regressing y on x and d^*_π is biased; see Goldberger (2008). A rough picture of these results is that Case 1 is easier to identify than Case 2. For the effect of measurement error in the average treatment effect framework, see Cochran and Rubin (1973), Rubin (1977) and Battistin and Chesher (2009).

A nonparametric framework rather than parametric framework is now commonly used in RDD analysis. Correspondingly, the *local polynomial estimator* (LPE) rather than the LSE is used in the estimation of $\alpha_\pi(\pi)$. In the nonparametric framework, we extend the results in the above literature in three aspects. First, for a fixed measurement error distribution, we re-analyze the identifiability of $\alpha_\pi(\pi)$ in the nonparametric model. Specifically, in Case 1, $\alpha_\pi(\pi)$ can be identified only in some special cases, e.g., when $\alpha_\pi(\cdot)$ is constant, or $\alpha_\pi(\cdot)$ is linear and the conditional distribution of u given $x = \pi$ is symmetric; in Cases 2 and 3, $\alpha_\pi(\pi)$ cannot be identified without further information. Second, by formulating a local measurement error that shrinks to zero with the sample size, we are able to sharpen the results in the first point. Specifically, in Case 1, $\alpha_\pi(\pi)$ can be identified, and compared to the Oracle Case, the LPE has a small extra bias but the same variance. In Case 2, $\alpha_\pi(\pi)$ can be identified, and compared with the Oracle Case, the LPE has a larger

[1] Such a measurement error is called a *non-differential* measurement error, see Carroll et al. (2006). There may also be a measurement error in y, but it should be absorbed in ε.

[2] The emphasis of the debate in the papers mentioned above lies in the sorting problem in heaping, that is, the nonrandom heaping. Without further clarification of the heaping mechanism, it is impossible to identify treatment effects in a nonrandom heaping.

bias and also a larger variance unless the measurement error is extremely small. In Case 3, $\alpha_\pi(\pi)$ cannot be identified unless the measurement error is extremely small. Third, we extend the analysis to the fuzzy design. Specifically, $\alpha_\pi(\pi)$ is easier to identify in Case 4 than in Case 5.

In passing, we mention that the usual measurement error literature assumes the distribution of u to be known from some validation data set, so a deconvolution estimator such as that in Fan and Truong (1993) can be used to recover the original conditional mean. Such a setup is not applicable to RDDs since data collection is usually conducted only once and no validation data set is available. In Case 2, Pei (2012) provides some interesting identification results by exploring the convolution relationship of $f_{x|D}$ with $f_{x*|D}$ and f_u, but one of the key assumptions in this approach is that the supports of x^* and u are discrete and bounded, which may restrict its applicability.[3]

The rest of the paper is organized as follows. In Section 2, we define the LPE with and without measurement error. When the measurement error is fixed, Section 3 and 4 analyze the identification of $\alpha_\pi(\pi)$ by the LPE when x is the forcing variable and when x^* is the forcing variable, respectively. Section 5 extends the analysis to the shrinking measurement error case. Section 6 turns to the analysis of the fuzzy design. Section 7 includes some simulation results and section 8 concludes. A word on notation: for a generic random variable y and a generic random vector x, $f_{x,y}(\cdot,\cdot)$ is their joint density, and $f_{y|x}(\cdot|\cdot)$ $(F_{y|x}(\cdot|\cdot))$ is the conditional pdf (cdf) of y given x; for a generic random variable x, $f_x(\cdot)$ $(F_x(\cdot))$ is the pdf (cdf) of x.

2 Existing Solutions in the Measurement Error Literature

Most solutions to the mismeasurement problem for the linear specification in econometrics depend on the use of instrumental variables. Another solution to the measurement error problem is to find an independent measure of the reliability of the variable, usually from a resampling approach. This solution can be given an instrumental variables interpretation. $Cov(z, x^*) \neq 0$ and $Cov(z, u) = Cov(z, \varepsilon_d) = 0$.

Reverse regression in Case 3.

Different from the literature, $d_i(\pi)$ is dummy, and is correlated with the measurement error in other regressors. The setup here is more or less close to that of Mahajan (2006) and Lewbel (2007), but in their case, other covariates are not mismeasured.

In Case 2, consider another estimator

$$\overleftrightarrow{\alpha}(\pi) = \overleftrightarrow{m}_+(\pi) - \overleftrightarrow{m}_-(\pi),$$

where $\overleftrightarrow{m}_+(\pi)$ is similarly defined as $\widehat{m}_+(\pi)$ but replacing $d_i(\pi)$ by $d_i^*(\pi)$, and in $\overleftrightarrow{m}_-(\pi)$, $1 - d_i(\pi)$ is replaced by $(1 - d_i^*(\pi))$. In other words, we use treated individuals in both neighborhoods of π to estimate $m_+^*(\pi)$, and controlled individuals in both neighborhoods of π to estimate $m_-^*(\pi)$.

Consider two identification schemes. First, when there are not further auxiliary information, we provide a partial identification result under the monotonicy assumption of Hausman, Abreveya, and Scott-Morton (1998). Second, when auxiliary information available, we provide point identification of the treatment effect and derive the asymptotic distribution of our estimator.

Hausman, Newey, Ichimura and Powell (1991, 1995) consider the polynomial setup. Ours is local polynomial, also there is a bianry regressor.

[3] Pei also discusses identifiability when x^* and u are continuously distributed with $u \sim N\left(0, \sigma^2\right)$. But this parametric assumption on the distribution of u is restrictive and may be questionable in practical work. For example, in Goldberger (2008)'s example, u must have a bounded support.

3 Local Polynomial Estimators

Suppose first that there is no measurement error in x^*. Since Porter (2003), the benchmark estimator of $\alpha_\pi(\pi)$ is the LPE. In the sharp design, it is defined as

$$\widehat{\alpha}^*(\pi) = \widehat{m}_+^*(\pi) - \widehat{m}_-^*(\pi), \tag{1}$$

where $\widehat{m}_+^*(\pi)$ is the LPE of $m_+^*(\pi) \equiv E[y|x^* = \pi+]$ and is determined by the minimizer $\widehat{a}$ in

$$\min_{a,b_1,\cdots,b_p} \frac{1}{n}\sum_{i=1}^n k_h(x_i^* - \pi) d_i^*(\pi)\left[y_i - a - b_1(x_i^* - \pi) - \cdots - b_p(x_i^* - \pi)^p\right]^2, \tag{2}$$

where p is a nonnegative integer, $d_i^*(\pi) = 1(x_i^* \geq \pi)$, $k_h(\cdot) = \frac{1}{h}k\left(\frac{\cdot}{h}\right)$, $k(\cdot)$ is a kernel function, and h is the bandwidth. $\widehat{m}_-^*(\pi)$ is the LPE of $m_-^*(\pi) \equiv E[y|x^* = \pi-]$ and is similarly defined as $\widehat{m}_+^*(\pi)$ but substituting $d_i^*(\pi)$ in (2) by $1 - d_i^*(\pi)$. The most popular choice of p is 1, and the resulting LPE is called the *local linear smoother* (LLS) which is popularized by Fan (1992, 1993). Under some regularity conditions, Porter (2003) shows that

$$\sqrt{nh}\left(\widehat{\alpha}^*(\pi) - \alpha_\pi(\pi) - B_n^*\right) \xrightarrow{d} N\left(0, \frac{\sigma_1^2(\pi) + \sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1\right), \tag{3}$$

where $\alpha_\pi(\pi) = m_+^*(\pi) - m_-^*(\pi)$, $B_n^* = h^{p+1}\frac{\alpha_\pi^{(p+1)}(\pi)}{(p+1)!}e_1'\Gamma_+^{-1}\mu_{p+1,2p+1}^+(1 + o_p(1))$ is a bias term, $e_1 = (1, 0, \cdots, 0)'_{(p+1)\times 1}$, $\sigma_d^2(x^*) \equiv E\left[\varepsilon_d^2|x^*\right]$, $d = 0, 1$, and

$$\Gamma_+ = \begin{pmatrix} \gamma_0^+ & \cdots & \gamma_p^+ \\ \vdots & \ddots & \vdots \\ \gamma_p^+ & \cdots & \gamma_{2p}^+ \end{pmatrix}_{(p+1)\times(p+1)}, \Omega_+ = \begin{pmatrix} \omega_0^+ & \cdots & \omega_p^+ \\ \vdots & \ddots & \vdots \\ \omega_p^+ & \cdots & \omega_{2p}^+ \end{pmatrix}_{(p+1)\times(p+1)}, \mu_{r,q}^+ = \begin{pmatrix} \gamma_r^+ \\ \vdots \\ \gamma_q^+ \end{pmatrix}, \tag{4}$$

with $\gamma_j^+ = \int_0^\infty k(u)u^j du$, $\omega_j^+ = \int_0^\infty k^2(u)u^j du$, and $r \leq q$ being nonnegative integers. In the fuzzy design, the LPE of $\alpha_\pi(\pi)$ is defined as

$$\widehat{\alpha}_f^*(\pi) = \frac{\widehat{m}_+^*(\pi) - \widehat{m}_-^*(\pi)}{\widehat{p}_+^*(\pi) - \widehat{p}_-^*(\pi)},$$

where $\widehat{p}_+^*(\pi)$ is the LPE of $p_+^*(\pi) \equiv E[D|x^* = \pi+]$ and is similarly defined as in (2) with D_i replacing y_i. $\widehat{p}_-^*(\pi)$ is similarly defined as $\widehat{p}_+^*(\pi)$ but using D_i in the left neighborhood of π. Under some regularity conditions, $\widehat{\alpha}_f^*(\pi)$ is consistent and has an asymptotically normal distribution.

When there is a measurement error u in x^*, only $x = x^* + u$ is observed. As mentioned in the introduction, there are two ways to assign treatment, based on either x or x^*. In the sharp design, if the forcing variable is x, then $D = d_\pi = 1(x \geq \pi)$ is always observed. The LPE of $\alpha_\pi(\pi)$ is

$$\widehat{\alpha}(\pi) = \widehat{m}_+(\pi) - \widehat{m}_-(\pi),$$

where $\widehat{m}_+(\pi)$ is the LPE of $m_+(\pi) \equiv E[y|x = \pi+]$, and is defined as the minimizer $\widehat{a}$ in

$$\min_{a,b_1,\cdots,b_p} \frac{1}{n}\sum_{i=1}^n k_h(x_i - \pi) d_i(\pi)\left[y_i - a - b_1(x_i - \pi) - \cdots - b_p(x_i - \pi)^p\right]^2, \tag{5}$$

where $d_i(\pi) = 1(x_i \geq \pi)$, and all other functions and variables are defined in (2). $\widehat{m}_-(\pi)$ is similarly defined as $\widehat{m}_+(\pi)$ but replacing $d_i(\pi)$ with $1 - d_i(\pi)$. If the forcing variable is x^*, we consider two cases. In Case

2, the treatment status $d_i^*(\pi)$ can be observed. In this case, it is natural to estimate $\alpha_\pi(\pi)$ by

$$\widetilde{\alpha}(\pi) = \widetilde{m}_+(\pi) - \widetilde{m}_-(\pi),$$

where $\widetilde{m}_+(\pi)$ is similarly defined as $\widehat{m}_+(\pi)$ but replacing $d_i(\pi)$ by $d_i(\pi)\, d_i^*(\pi)$, and in $\widetilde{m}_-(\pi)$, $1 - d_i(\pi)$ is replaced by $(1 - d_i(\pi))(1 - d_i^*(\pi))$. In other words, we use treated individuals in the right neighborhood of π to estimate $m_+^*(\pi)$, and controlled individuals in the left neighborhood of π to estimate $m_-^*(\pi)$. In Case 3, $d_i^*(\pi)$ cannot be observed, then we have to estimate $\alpha_\pi(\pi)$ by $\widehat{\alpha}(\pi)$.

In the fuzzy design, we assume the treatment status can be observed, otherwise the causal effect cannot be identified even if x^* is not contaminated. If the forcing variable is x, assume $D = p_\pi(x) + \beta_\pi(x) d_\pi + \eta$, where $E[\eta|x] = 0$, $p_\pi(x) \neq 0$ and $\beta_\pi(\pi) \neq 0$. Now, the LPE of $\alpha_\pi(\pi)$ is

$$\widehat{\alpha}_f(\pi) = \frac{\widehat{m}_+(\pi) - \widehat{m}_-(\pi)}{\widehat{p}_+(\pi) - \widehat{p}_-(\pi)}, \tag{6}$$

where $\widehat{p}_+(\pi)$ ($\widehat{p}_-(\pi)$) is an estimator of $p_+(\pi) \equiv E[D|x = \pi+]$ ($p_-(\pi) \equiv E[D|x = \pi-]$) and is similarly defined as in (5) but replacing y_i by D_i $(1 - D_i)$. When x^* is the forcing variable, as in the case without measurement error, $D = p_\pi(x^*) + \beta_\pi(x^*) d_\pi^* + \eta$. We still use $\widehat{\alpha}_f(\pi)$ to estimate $\alpha_\pi(\pi)$.

The goal of this paper is to check whether the LPE can identify $\alpha_\pi(\pi)$ in different cases. For future reference, we define some basic density or conditional density functions and state some properties of them. Suppose the joint density of (x^*, ε_d) is $f^d_{x^*,\varepsilon}(x^*, \varepsilon)$, then the joint density of (x, ε_d, u) is $f^d_{x^*,\varepsilon}(x - u, \varepsilon) f_u(u)$, and the joint density of (x, ε_d) is $\int f^d_{x^*,\varepsilon}(x - u, \varepsilon) f_u(u) du$, $d = 0, 1$. Also, the joint density of (x, u) is $f_{x^*}(x - u) f_u(u)$ and the marginal density of x is $f_x(x) = \int f_{x^*}(x - u) f_u(u) du$. Note that

$$\begin{aligned} E[\varepsilon_d|x] &= \int \varepsilon \frac{\int f^d_{x^*,\varepsilon}(x-u,\varepsilon) f_u(u) du}{\int f_{x^*}(x-u) f_u(u) du} d\varepsilon = \frac{\int \int \varepsilon f^d_{x^*,\varepsilon}(x-u,\varepsilon) d\varepsilon f_u(u) du}{\int f_{x^*}(x-u) f_u(u) du} \\ &= \int E[\varepsilon_d|x^* = x-u] \frac{f_{x^*}(x-u) f_u(u)}{\int f_{x^*}(x-u) f_u(u) du} du = \int E[\varepsilon_d|x^* = x-u] f_{u|x}(u|x) du = 0, \end{aligned}$$

where the second equality is from Fubini's theorem, and the last equality is from the assumption that $E[\varepsilon_d|x^*] = 0$.

In the measurement error literature, it usually assumes that $E[\varepsilon_d|x, x^*] = E[\varepsilon_d|x^*] = 0$ (referred as the assumption of *nondifferential measurement error* by Mahajan (2006)), which implies $E[\varepsilon_d|x] = 0$. This assumption excludes the placebo effect and implies that the measurement error is not driven by the unobserved benefit. Bound, Brown, and Mathiowetz (2000) provide instances when such an assumption is likely or unlikely to hold; see page 540 of Lewbel (2007) for futher discussions.

$$P(\pi + h \geq x \geq \pi | x^* < \pi) = \frac{\int_{-\infty}^{\pi} \int_{\pi}^{\pi+h} f_u(x - x^*) dx f_{x^*}(x^*) dx^*}{\int_{-\infty}^{\pi} f_{x^*}(x^*) dx^*}$$

4 The Forcing Variable is x

If the treatment is based on x, $y = m_\pi(x^*) + \alpha_\pi(x^*)\, d_\pi + \varepsilon$. In the LPE $\widehat{\alpha}(\pi)$, $\widehat{m}_+(\pi)$ and $\widehat{m}_-(\pi)$ are estimating $m_+(\pi)$ and $m_-(\pi)$, respectively. The question is what $m_+(\pi) - m_-(\pi)$ is identifying. As discussed at the end of section 2, $E[\varepsilon|x] = 0$, so we need only concentrate on $E[m_\pi(x^*)|x]$ and $E[\alpha_\pi(x^*)|x]$.

Note that

$$E[m_\pi(x^*)|x] = E[m_\pi(x - u)|x] = \int m_\pi(x - u) f_{u|x}(u|x) du \equiv \overline{m}_\pi(x)$$

is continuous if both $m_\pi(\cdot)$ and $f_{u|x}(u|\cdot)$ are continuous. So if we define

$$\overline{\alpha}_\pi(x) = E[\alpha_\pi(x^*)|x] = \int \alpha_\pi(x-u) f_{u|x}(u|x)du, \tag{7}$$

then

$$\begin{aligned} & m_+(\pi) - m_-(\pi) \\ = \ & E[m_\pi(x^*)|x=\pi+] - E[m_\pi(x^*)|x=\pi-] + E\left[\alpha_\pi\left(x^*\right)|x=\pi+\right] \\ = \ & \int \alpha_\pi(\pi-u) f_{u|x}(u|\pi)du = \overline{\alpha}_\pi(\pi) \\ = \ & \int \alpha_\pi(\pi-u)\frac{f_{x|u}(\pi|u)}{f_x(\pi)} f(u)du = E_u\left[\alpha_\pi(\pi-u)\frac{f_{x|u}(\pi|u)}{f_x(\pi)}\right]. \end{aligned}$$

where the second equality is from the continuity of $\alpha_\pi(\cdot)$ and $f_{u|x}(u|\cdot)$. So $m_+(\pi) - m_-(\pi)$ is a weighted average treatment effect with weight $\frac{f_{x|u}(\pi|u)}{f_x(\pi)}$ which is directly proportional to the likelihood that the realization of x corresponding to u will be close to the threshold π. If we interpret u as the unobserved component of the outcome in Lee (2008), then from his Proposition 3, $\alpha_\pi(\pi-u)$ plays the role of $y(\pi+,u) - y(\pi-,u)$, where $y(x,u)$ defines the relationship between y and (x,u).

We are interested in the cases where $\overline{\alpha}_\pi(\pi) = \alpha_\pi(\pi)$. Suppose $\alpha_\pi(x^*)$ is smooth enough, then $\alpha_\pi(\pi - u) \approx \alpha_\pi(\pi) - \alpha'_\pi(\pi)u + \frac{1}{2}\alpha''_\pi(\pi)u^2$ to the second-order polynomial. If both $\alpha'_\pi(\pi)$ and $\alpha''_\pi(\pi)$ are zero, and $\alpha_\pi(\pi-u) = \alpha_\pi(\pi)$, then $\overline{\alpha}_\pi(\pi) = \alpha_\pi(\pi)$; that is, in the common treatment effect case, $\alpha_\pi(\pi)$ can be identified. Cappelleri et al. (1991) and Goldberger (2008) show this result in a very special case as mentioned in the introduction. Here, we show this result in the general setup where $m_\pi(x^*)$ is not necessarily linear and the distributions of x^* and u are not necessarily normal. If $\alpha'_\pi(\pi) \neq 0$, $\alpha''_\pi(\pi) = 0$, and $f_{u|x}(u|\pi)$ is symmetric about zero, then $\overline{\alpha}_\pi(\pi)$ is also equal to $\alpha_\pi(\pi)$. In these two cases, although $E[y|x]$ may be different from $m(x)$, the jump size at π remains. If $\alpha''_\pi(\pi) \neq 0$, because $E[u^2|x=\pi] > 0$, it is hard to imagine when $\overline{\alpha}_\pi(\pi)$ will equal $\alpha_\pi(\pi)$. The following two examples provides more intuition on the identification of $\alpha_\pi(\pi)$.

Example 1: Suppose $\alpha_\pi(x^*) = \alpha_0 + \alpha_1 x^*$, then

$$\overline{\alpha}_\pi(x) = \int \left(\alpha_0 + \alpha_1(x-u)\right) f_{u|x}(u|x)du = \alpha_0 + \alpha_1 x - \alpha_1 E[u|x],$$

where $-\alpha_1 E[u|x]$ is an extra term introduced by measurement error, which is equal to zero iff $\alpha_1 = 0$ or $E[u|x] = 0$. If $E[u|x]$ takes a linear form such as in the case where both u and x^* are normally distributed, then $E[u|x]$ is the linear projection of u on x, and $E[u|x] = -(1-\rho)E[x] + (1-\rho)x$, where $\rho = \frac{Var(x^*)}{Var(x^*)+Var(u)}$ is the *reliability coefficient*. In consequence,

$$\overline{\alpha}_\pi(x) = \alpha_0 + \alpha_1(1-\rho)E[x] + \rho\alpha_1 x \equiv \widetilde{\alpha}_0 + \widetilde{\alpha}_1 x,$$

where $E[x] = E[x^*]$ when $E[u] = 0$. The estimate of the *interaction effect* α_1 is biased by a factor ρ.[4] Similarly, if $m_\pi(x^*) = \beta_0 + \beta_1 x^*$, and $E[u|x]$ is linear in x,

$$\overline{m}_\pi(x) = \beta_0 + \beta_1(1-\rho)E[x] + \rho\beta_1 x \equiv \widetilde{\beta}_0 + \widetilde{\beta}_1 x.$$

[4] The term "interaction effect" is used in the literature, e.g., Trochim et al. (1991), to refer the fact that the treatment effect is a function of x^* rather than a constant.

As in the usual analysis of measurement error, the attenuation effect appears: $\left|\widetilde{\beta}_1\right| \leq |\beta_1|$. In summary,

$$\begin{aligned} E[y|x] &= \widetilde{\beta}_0 + \widetilde{\beta}_1 x + (\widetilde{\alpha}_0 + \widetilde{\alpha}_1 x)\, d_\pi \\ &= \left(\widetilde{\beta}_0 + \widetilde{\beta}_1 x\right) 1(x < \pi) + \left[\left(\widetilde{\beta}_0 + \widetilde{\alpha}_0\right) + \left(\widetilde{\beta}_1 + \widetilde{\alpha}_1\right) x\right] 1(x \geq \pi). \end{aligned}$$

If we regress y on x, d_π and xd_π, we can recover $\widetilde{\beta}_0$, $\widetilde{\beta}_1$, $\widetilde{\alpha}_0$, and $\widetilde{\alpha}_1$. If ρ were known, then α_0 and α_1 can be recovered from the spurious regression in an obvious way:

$$\alpha_1 = \frac{\widetilde{\alpha}_1}{\rho}, \alpha_0 = \widetilde{\alpha}_0 - \alpha_1(1-\rho)E[x] = \widetilde{\alpha}_0 - \widetilde{\alpha}_1 \frac{1-\rho}{\rho} E[x].$$ [5]

In the general case, $m_\pi(\cdot)$ need not be linear, so the usual linear regression of y on x and d_π can introduce bias to $\alpha_\pi(\pi)$ even if $\alpha_\pi(\cdot)$ is constant. Trochim et al. (1991) noticed the harm of misspecification of $m_\pi(\cdot)$, but did not dig further. Because nonparametric techniques are usually used in the RD analysis, the misspecification problem in $m_\pi(\cdot)$ can be successfully conquered.

The left two panels of Figure 1 show $E[y|x]$ and $E[y|x^*]$ when $\alpha_1 = 0$ and $\alpha_1 \neq 0$, respectively. In both panels, x^* and u are normally distributed. $x^* \sim N(2, 1)$ and $u \sim N\left(0, \frac{3}{7}\right)$, so the reliability coefficient $\rho = 0.7$. In the common treatment effect case, $\alpha_0 = \beta_1 = 1$, and $\alpha_1 = \beta_0 = 0$; in the variable treatment effect case, $\alpha_0 = \beta_0 = 0$, and $\alpha_1 = \beta_1 = 1$. In both cases, $\pi = 1$, and $\alpha_\pi(\pi) = 1$. In the upper-left panel, although $E[y|x]$ gets flatter than $E[y|x^*]$, the jump sizes at π are the same. In the lower-left panel, the jump sizes at π may not be the same, but the jump size of $E[y|x]$ at π, $\overline{\alpha}_\pi(\pi)$, is equal to $\alpha_\pi(\cdot)$ evaluated at a different point $\overline{\pi}$. It is easy to show that $\overline{\alpha}_\pi(\pi) = \alpha_\pi(\overline{\pi})$ with $\overline{\pi} = E[x] - \rho(E[x] - \pi)$, which is equivalent to the formula (4) in Trochim et al. (1991). This means that the estimated treatment effect at π is the true treatment effect at

$$\overline{\pi} = E[x] - \rho(E[x] - \pi).$$

Figure 2 shows why $\overline{\pi} \neq \pi$. When $\rho = 1$, that is, there is no measurement error, $E[x] - \rho(E[x] - \pi) = \pi$. Of course, when $\rho = 1$, $u = 0$ almost surely, and $\overline{\alpha}_\pi(x) = \int \alpha_\pi(x-u) f_{u|x}(u|x) du = \alpha_\pi(x)$ even in the general case; see more analysis in Section 5.1 below. Also, when $\pi = E[x]$, $E[u|x = \pi] = 0$, so $\overline{\pi} = \pi$. $\square$

Example 1 shows that when there is a measurement error in x^* ($\rho < 1$) and $\alpha_\pi(x^*)$ is linear and nonconstant, $\overline{\alpha}_\pi(\pi)$ is not equal to $\alpha_\pi(\pi)$ if $E[u|x]$ is linear unless $\pi = E[x]$. If $E[u|x]$ is nonlinear, the situation is more complicated.

Example 2: Suppose $\alpha_\pi(x^*) = \alpha_0 + \alpha_1 x^*$ and $x^* \sim U[0, 3]$. The two upper panels of Figure 3 shows the support of (x, u) and $E[u|x]$ when $u \sim U[-1, 1]$. Since $\overline{\alpha}_\pi(x) = \alpha_0 + \alpha_1 x - \alpha_1 E[u|x]$, it is straightforward to see that $\overline{\alpha}_\pi(\pi) = \alpha_\pi(\pi)$ when $\pi \in [1, 2]$ even if $\alpha_1 \neq 0$. The two lower panels of Figure 3 shows the support of (x, u) and $E[u|x]$ when $u \sim N(0, 1)$. In this case, only if $\pi = 1.5 = E[x]$, $\overline{\alpha}_\pi(\pi) = \alpha_\pi(\pi)$ when $\alpha_1 \neq 0$. All relevant calculations are included in the supplementary materials. $\square$

The theorem below rigorously states the asymptotic distribution of the LPE $\widehat{\alpha}(\pi)$. First, we specify some regularity conditions.

Assumption K: $k(\cdot)$ is a symmetric, bounded, Lipschitz function, zero outside a bounded set $[-M, M]$, and $\int k(u) du = 1$.

[5] Cappelleri et al. (1991) provide similar formulas for β_0 and β_1, but the corrected intercept estimator in their equation (5) is wrong because β_1 is lost. Since $\beta_1 = 1$ in their simulations, β_0 is not affected and their simulation results are still valid.

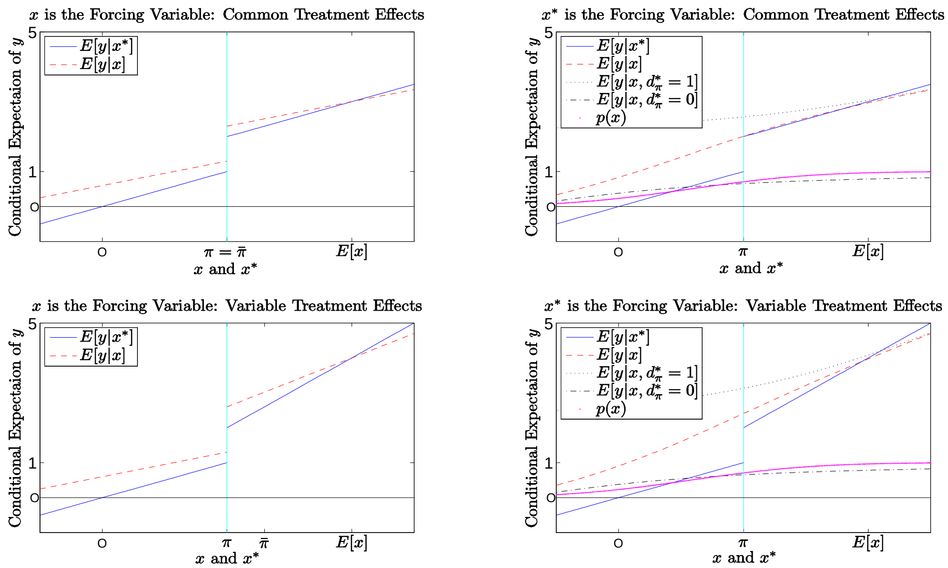

Figure 1: $E[y|x^*]$, $E[y|x]$, $E[y|d_\pi^* = 1, x]$ and $E[y|d_\pi^* = 0, x]$ When x or x^* is the Forcing Variable: $\rho = 0.7$

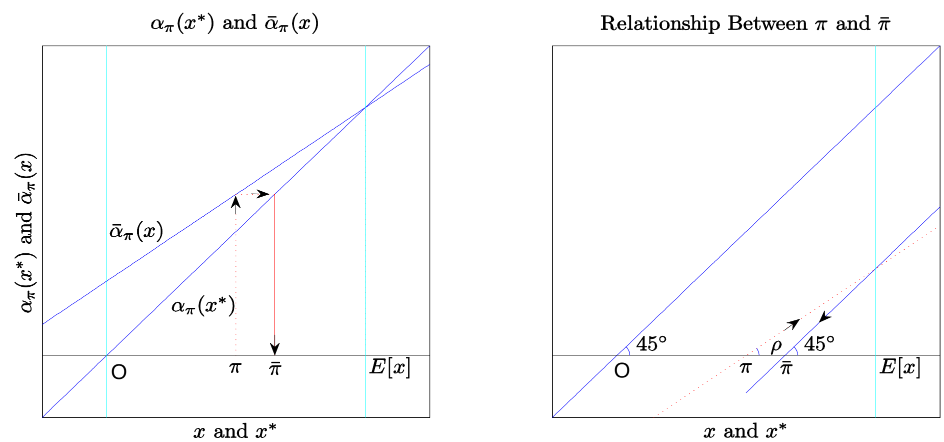

Figure 2: Illustration of $\bar{\pi} \neq \pi$ When $\alpha_\pi(x^*)$ is not Constant, $\pi \neq E[x]$ and $\rho \neq 1$

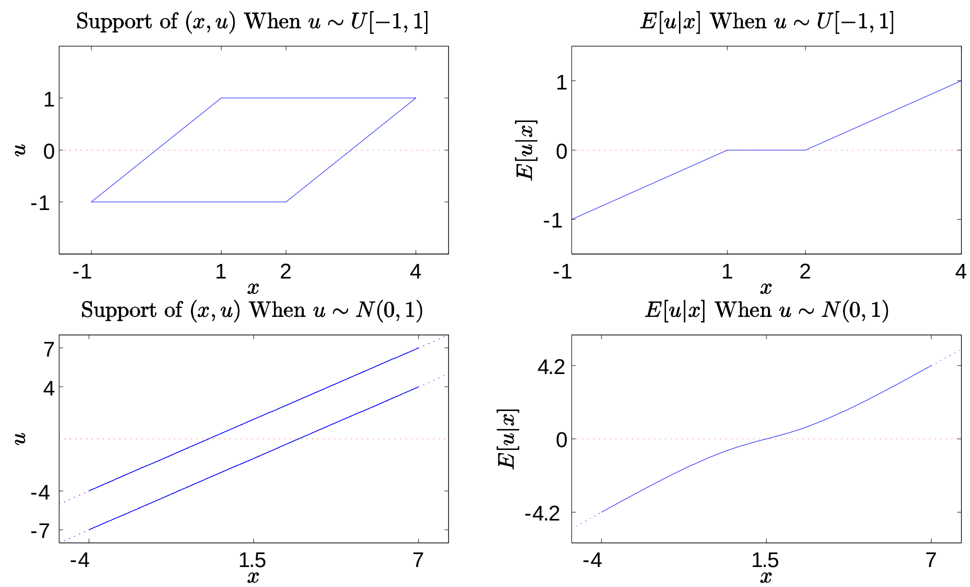

Figure 3: Support of (x,u) and $E[u|x]$ when $u \sim U[-1,1]$ or $u \sim N(0,1)$: $x^* \sim U[0,3]$

We assume $k(\cdot)$ to have a compact support to simplify the proof. In the discussion below, we let $M = 1$ without loss of generality.

Assumption F: $f_{x^*}(x^*)$ and $f_u(u)$ are continuous and uniformly bounded on their supports $\mathcal{X}^*$ and $\mathcal{U}$, respectively, and $f_x(\pi)$ is bounded away from 0.

Since $f_x(x) = \int f_{x^*}(x-u) f_u(u) du$, Assumption F implies that $f_x(x)$ is continuous and uniformly bounded on its support.

Assumption M: $m_\pi(x^*)$ and $\alpha_\pi(x^*)$ are continuous and uniformly bounded on $\mathcal{X}^*$. $\alpha_\pi(\pi) \neq 0$.

Assumption EY: For $d = 0, 1$,

(a) $E[\varepsilon_d | x^*] = 0$.

(b) $\sigma_d^2(x^*) = E\left[\varepsilon_d^2 | x^*\right]$ is continuous and uniformly bounded on $\mathcal{X}^*$.

(c) For some $\zeta > 0$, $E\left[\left|\varepsilon_d\right|^{2+\zeta} \middle| x^*\right]$ is uniformly bounded on $\mathcal{X}^*$.

Assumption B: $h \to 0$ and $nh \to \infty$.

All these assumptions are quite standard. The boundedness assumptions in EY, F and M are extended to the whole support of x^* instead of a neighborhood of π as in Porter (2003). This is obviously because the measurement error contaminates the distribution of x^*.

Theorem 1 *Under Assumptions B, EY, F, K and M, in Case 1,*

$$\sqrt{nh}\left(\widehat{\alpha}(\pi) - \overline{\alpha}_\pi(\pi) - B_{1n}\right) \xrightarrow{d} N\left(0, \frac{E\left[\varepsilon_1^2 | x = \pi\right] + E\left[\varepsilon_0^2 | x = \pi\right]}{f_x(\pi)} e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1\right),$$

where $\overline{\alpha}_\pi(\pi)$ *is defined in (7),* $B_{1n} = o_p(1)$ *is defined in (16), and* Γ_+ *and* Ω_+ *are defined in (4).*

A straightforward corollary of Theorem 1 is that $\widehat{\alpha}(\pi)$ is a consistent estimator of $\overline{\alpha}_\pi(\pi)$ which may not be equal to $\alpha_\pi(\pi)$ except in some special cases. The asymptotic variance is similar as in the no-measurement-error case (3), but in each term involving x^*, x^* is replaced by x.

5 The Forcing Variable is x^*

In Case 2, we use $\widetilde{\alpha}(\pi)$ to estimate $\alpha_\pi(\pi)$. $\widetilde{\alpha}(\pi)$ is estimating $E[y|d^*_\pi = 1, x = \pi+] - E[y|d^*_\pi = 0, x = \pi-]$, so we need only check whether it is equal to $\alpha_\pi(\pi)$ or not. In Case 3, we use $\widehat{\alpha}(\pi)$ to estimate $\alpha_\pi(\pi)$, where we check whether $m_+(\pi) - m_-(\pi)$ is equal to $\alpha_\pi(\pi)$ or not. Note that if the forcing variable is x^*, $m(x^*)$ can always be written as $m^*_\pi(x^*) + \alpha_\pi(\pi)d^*_\pi$, where $m^*_\pi(x^*)$ is a continuous function defined as $m(x^*) - \alpha_\pi(\pi)\,d^*_\pi$. It is easy to see that $m^*_\pi(x^*) = m_\pi(x^*)$ when $x^* < \pi$, and $m^*_\pi(x^*) = m_\pi(x^*) + \alpha_\pi(x^*) - \alpha_\pi(\pi)$ when $x^* \geq \pi$.

In the supplementary materials, we show that

$$\begin{aligned}
&E[y|d^*_\pi = 1, x] - E[y|d^*_\pi = 0, x] \\
&= \alpha_\pi(\pi) + E[m^*_\pi(x^*)|d^*_\pi = 1, x] - E[m^*_\pi(x^*)|d^*_\pi = 0, x] \\
&= \alpha_\pi(\pi) + \frac{\int_\pi^\infty m^*_\pi(x^*) f_{x^*|x}(x^*|x)dx^*}{1 - F_{x^*|x}(\pi|x)} - \frac{\int_{-\infty}^\pi m^*_\pi(x^*) f_{x^*|x}(x^*|x)dx^*}{F_{x^*|x}(\pi|x)} \\
&\equiv \alpha_\pi(\pi) + \overline{m}^+_\pi(x) - \overline{m}^-_\pi(x) \equiv \alpha_\pi(\pi) + \Delta\overline{m}_\pi(x).
\end{aligned} \tag{8}$$

where the first equality is from $E[\varepsilon|d^*_\pi = 1, x] = E[\varepsilon|d^*_\pi = 0, x] = 0$. So we must have $\Delta\overline{m}_\pi(\pi) = 0$ to identify $\alpha_\pi(\pi)$. Note that

$$\Delta\overline{m}_\pi(\pi) = \frac{\int_{-\infty}^0 m^*_\pi(\pi - u) f_{x^*|x}(\pi - u|\pi)\, du}{\int_{-\infty}^0 f_{x^*|x}(\pi - u|\pi)\, du} - \frac{\int_0^\infty m^*_\pi(\pi - u) f_{x^*|x}(\pi - u|\pi)\, du}{\int_0^\infty f_{x^*|x}(\pi - u|\pi)\, du}$$

is the difference between the conditional mean of $m^*_\pi(x^*)$ given $x = \pi$ with $u < 0$ and $u > 0$, so $\Delta\overline{m}_\pi(\pi)$ is generally not zero unless $m^*_\pi(\cdot)$ is constant. In Case 3,

$$\begin{aligned}
E[y|x] &= E[m(x^*)|x] = E[m^*_\pi(x^*)|x] + \alpha_\pi(\pi)\, p(x) \\
&= E[m(x^*)|d^*_\pi = 0, x]\, P(d^*_\pi = 0|x) + E[m(x^*)|d^*_\pi = 1, x,]\, P(d^*_\pi = 1|x) \\
&= E[m(x^*)|d^*_\pi = 0, x]\,(1 - p(x)) + E[m(x^*)|d^*_\pi = 1, x]\, p(x)
\end{aligned} \tag{9}$$

which is an average of $E[m(x^*)|d^*_\pi = 0, x]$ and $E[m(x^*)|d^*_\pi = 1, x]$, where

$$p(x) = E[d^*_\pi|x] = P(x - u \geq \pi|x) = F_{u|x}(x - \pi)$$

is the propensity score. Because $E[m^*_\pi(x^*)|x]$ and $p(x)$ are both continuous, $E[y|x]$ is continuous. Therefore, $m_+(\pi) - m_-(\pi)$ is always 0 and cannot identify $\alpha_\pi(\pi)$. Another observation is that this is a fuzzy design in terms of Trochim (1984), so we can check whether $\frac{m_+(\pi) - m_-(\pi)}{p_+(\pi) - p_-(\pi)} = \alpha_\pi(\pi)$, where $p_\pm(\pi) = E[d^*_\pi = 1|x = \pi\pm]$. But both the numerator and denominator are zero since $E[y|x]$ and $p(x)$ are both continuous, so $\frac{m_+(\pi) - m_-(\pi)}{p_+(\pi) - p_-(\pi)}$ cannot equal $\alpha_\pi(\pi)$. The following example, which uses a similar setup as in Example 1, illustrates the above points intuitively.

Example 3: Suppose $x^* \sim N(\mu, \sigma^2)$, and $u \sim N\left(0, \frac{1-\rho}{\rho}\sigma^2\right)$, then the reliability coefficient is ρ. Suppose further that $\alpha_\pi(x^*) = \alpha_0 + \alpha_1 x^*$, and $m_\pi(x^*) = \beta_0 + \beta_1 x^*$, then $\alpha_\pi(\pi) = \alpha_0 + \alpha_1 \pi$ and

$$m_\pi^*(x^*) = (\beta_0 + \beta_1 x^*) 1(x^* < \pi) + (\overline{\beta}_0 + \overline{\beta}_1 x^*) 1(x^* \geq \pi),$$

where $\overline{\beta}_0 = \beta_0 - \alpha_1 \pi$, and $\overline{\beta}_1 = \alpha_1 + \beta_1$. In the supplementary materials, we extend the analysis in Appendix C of Goldberger (2008) to show that

$$\begin{aligned} \overline{m}_\pi^+(x) &= \overline{\beta}_0 + \overline{\beta}_1 \left[a(x) + b\lambda\left(\frac{a(x)-\pi}{b}\right)\right], \\ \overline{m}_\pi^-(x) &= \beta_0 + \beta_1 \left[a(x) - b\lambda\left(\frac{\pi - a(x)}{b}\right)\right], \end{aligned}$$

where $a(x) = \rho x + (1-\rho)\mu$, $b = \sqrt{1-\rho}\sigma$, and $\lambda(z) = \frac{\phi(z)}{\Phi(z)}$ is the *inverse Mill's ratio* with $\phi(\cdot)$ and $\Phi(\cdot)$ being the pdf and cdf of the standard normal distribution, respectively. So

$$\Delta\overline{m}_\pi(\pi) = \alpha_1 \left[\Delta + b\lambda\left(\frac{\Delta}{b}\right)\right] + \beta_1 b \left[\lambda\left(\frac{\Delta}{b}\right) - \lambda\left(-\frac{\Delta}{b}\right)\right]$$

where $\Delta = (1-\rho)(\mu - \pi)$. When $\Delta \neq 0$ (that is, $\rho < 1$ and $\pi \neq \mu$), $\Delta\overline{m}_\pi(\pi) \neq 0$ unless $\alpha_1 = 0$ and $\beta_1 = 0$ which corresponds to a constant $m_\pi^*(\cdot)$. But a constant $m_\pi^*(\cdot)$ rarely happens in practice; see, e.g., Figure 2-5 of Lee (2008) to check this fact. Furthermore,

$$E[y|x] = [\beta_0 + \beta_1 a(x)] + \alpha_1 (a(x) - \pi)\Phi\left(\frac{a(x)-\pi}{b}\right) + \alpha_1 b\phi\left(\frac{a(x)-\pi}{b}\right) + \alpha_\pi(\pi) p(x),$$

where $p(x) = \Phi\left(\frac{a(x)-\pi}{b}\right)$. The right two panels of Figure 1 show $E[y|x^*]$, $E[y|x]$, $E[y|d_\pi^* = 1, x]$, $E[y|d_\pi^* = 0, x]$ and $p(x)$ for the common and variable treatment effect case, respectively, where all relevant parameter values are the same as in Example 1. It is obvious that neither $E[y|d_\pi^* = 1, x = \pi+] - E[y|d_\pi^* = 0, x = \pi-]$ nor $E[y|x = \pi+] - E[y|x = \pi-]$ can identify $\alpha_\pi(\pi)$ even in the common treatment effect case. In this parametric model, nonlinear least squares or maximum likelihood can be used to identify $\alpha_\pi(\pi)$, but they are generally inapplicable in nonparametric settings. □

The following theorem rigorously states the probability limit of $\widetilde{\alpha}(\pi)$ and $\widehat{\alpha}(\pi)$.

Theorem 2 *Under Assumptions B, EY, F, K and M, in Case 2,*

$$\widetilde{\alpha}(\pi) \xrightarrow{p} \alpha_\pi(\pi) + \Delta\overline{m}_\pi(\pi),$$

and in Case 3,

$$\widehat{\alpha}(\pi) \xrightarrow{p} 0,$$

where $\Delta\overline{m}_\pi(\cdot)$ is defined in (8).

Comparing with Theorem 1, the bias of $\widetilde{\alpha}(\pi)$ is also affected by $m_\pi(\cdot)$ besides $\alpha_\pi(\cdot)$, while the bias of $\widehat{\alpha}(\pi)$ in Case 1 is only affected by $\alpha_\pi(\cdot)$.

5.1 Unidentifiability of the Treatment Effect

When the forcing variable is x^*, $\alpha_\pi(\pi)$ cannot be identified. To see why, rewrite the model in familiar notations as

$$Y(D) = \mu_D(x) + U_D \text{ with } D = 1(x - \pi - u \geq 0), \tag{10}$$

where

$$\begin{aligned}
\mu_0(x) &= E\left[m_\pi^*(x-u)|x, D=0\right], \\
\mu_1(x) &= E\left[m_\pi^*(x-u)|x, D=1\right] + \alpha_\pi(\pi), \\
U_0 &= \varepsilon_0 + m_\pi^*(x-u) - E\left[m_\pi^*(x-u)|x, D=0\right], \\
U_1 &= \varepsilon_1 + m_\pi^*(x-u) - E\left[m_\pi^*(x-u)|x, D=1\right].
\end{aligned}$$

Since both U_0 and U_1 include the component of u, they are correlated with D given x, so the unconfoundedness condition fails, and $\alpha_\pi(\pi)$ cannot be identified. Actually, this is a model of *essential heterogeneity* in terms of Heckman and Vytlacil (see, e.g., Heckman et al. (2006)), since

$$U_1 - U_0 = \varepsilon_1 - \varepsilon_0 + E\left[m_\pi^*(x-u)|x, u > x-\pi\right] - E\left[m_\pi^*(x-u)|x, u \le x-\pi\right]$$

is correlated with D after controlling for x. Of course, there is also a *selection bias* because D is also correlated with U_0. This unidentifiability result implies that the identification technique in Pei (2011) critically relies on the discreteness of x^* and u. His identification scheme cannot be extended to the general case where the distribution of x^* and u are continuous and nonparametrically unknown.[6]

For comparison, we review how a benchmark treatment model is specified. Suppose we still use the threshold crossing model to assign the treatment status, then D should be equal to $1(x^* - \pi + u \ge 0)$, and x^* is observed. In this case,

$$\begin{aligned}
Y(D) &= m_\pi(x^*) + \alpha_\pi(x^*) D + U_D \text{ with } U_D = \varepsilon_D, \\
D &= 1(x^* - \pi + u \ge 0).
\end{aligned} \tag{11}$$

Now, $(Y(0), Y(1))$ are conditional independent of D given x^*, since we assume that u is independent of $(\varepsilon_0, \varepsilon_1)$. The average treatment effect in this ideal case is $E[\alpha_\pi(x^*)]$. In Case 1,

$$\begin{aligned}
Y(D) &= E\left[m_\pi(x-u)|x\right] + E\left[\alpha_\pi(x-u)|x\right] D + U_D, \\
D &= 1(x - \pi \ge 0),
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
U_0 &= \varepsilon_0 + m_\pi(x-u) - E\left[m_\pi(x-u)|x\right], \\
U_1 &= \varepsilon_1 + m_\pi(x-u) + \alpha_\pi(x-u) - E\left[m_\pi(x-u) + \alpha_\pi(x-u)|x\right].
\end{aligned}$$

Conditional on x, D is a constant, so the unconfoundedness condition is satisfied. The difficulty here is that the overlap assumption is violated because given any x, we can observe either $Y(0)$ or $Y(1)$ but not both. Of course, we can use the continuity of $E\left[m_\pi(x-u)|x\right]$ and $E\left[\alpha_\pi(x-u)|x\right]$ to extend $E[Y|x]$ to the opposite side of π, but the measured treatment effect $E\left[\alpha_\pi(x-u)|x=\pi\right]$ may not be equal to the true treatment effect $\alpha_\pi(\pi)$ unless in the neighborhood of $x=\pi$, $\alpha_\pi(x-u) \approx \alpha_\pi(x) + g(u)$ with $E[g(u)|x] = 0$; see Section 3 for some special cases such that this is satisfied. The setup of Case 2, (10), deviates from the benchmark (11) by assuming that the observed assignment covariate x is correlated with the assignment error u such that the unconfoundedness condition fails. This is a more severe deviation from the benchmark model. In Case 3, even the treatment status cannot be observed. In his postcript, Goldberger (2008) says that his

[6] Hullegie and Klein (2010) consider identification of the treatment effects in a parametric model where u is independent of x (rather than x^*) and follows $N(0, \sigma_u^2)$.

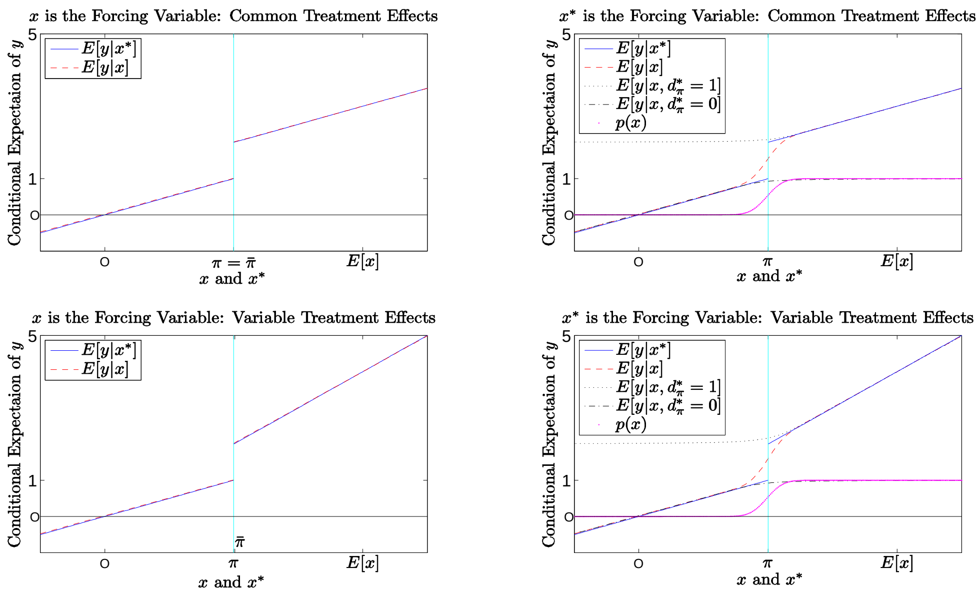

Figure 4: $E[y|x^*]$, $E[y|x]$, $E[y|d^*_\pi = 1, x]$ and $E[y|d^*_\pi = 0, x]$ When x^* or x is the Forcing Variable: $\rho = 0.99$

parametric analysis about Case 1 and Case 2 "is best viewed as making the distinction between selection on observables and selection on unobservables". His comments obviously apply to the nonparametric setup of this paper.

A common solution in Case 2 to identify $\alpha_\pi(\pi)$ is to use an instrument variable that introduces extra randomness in D and is independent of U_D. For example, let the instrument variable be correlated with x^* and independent of u, ε_0 and ε_1. But this is against RDDs at the beginning since the key point of RDDs is to avoid using the instrument variable to control endogeneity and assign the treatment based on a single forcing variable.

6 Shrinking Measurement Error

When $m_\pi(\cdot)$ and $\alpha_\pi(\cdot)$ take a parametric form, the unit of the measurement error should be the usual unit and similar to the unit of x^*. Just like in the linear regression, the unit of the error term is the same as the regressors. In the nonparametric setup, $m_\pi(\cdot)$ and $\alpha_\pi(\cdot)$ take a parametric form only in a h neighborhood of each point on the support of x^*, so it is natural to use h as the unit of the measurement error. In other words, we should assume the measurement error shrinking to zero to check the identifiability of the treatment effect.[7] Figure 4 shows the same information as in Figure 1 except that $\rho = 0.99$ now. From Figure 4, it seems that Case 1 is easier than Case 2 which is in turn easier than Case 3 in identification of the treatment effect. The discussion below tries to strengthen this intuition by letting the measurement error shrink to zero. To simplify our analysis, suppose $u = n^{-\delta}\epsilon$ for some random variable ϵ and $\delta > 0$. This form of measurement error is easy to compare with h. For example, the usual bandwidth $h = \varrho n^{-1/5}$, $\varrho > 0$, so we can compare δ with $1/5$ to check the limit of $\frac{n^{-\delta}}{h}$ which will be used in the following theorems.

[7] Battistin and Chesher (2009) also assume the measure error shrinking to zero in the average treatment effect framework under the uncounfoundedness assumption.

6.1 The Forcing Variable is x

Before stating the asymptotic distribution of the LPE, we adapt Assumption F to this new environment.

Assumption F′: $u = n^{-\delta}\epsilon$, $\delta > 0$. $f_{x^*}(x^*)$ and $f_\epsilon(\epsilon)$ are continuous and uniformly bounded on their supports, and $f_{x^*}(\pi)$ is bounded away from 0.

Theorem 3 *Under Assumptions B, EY, F′, K and M, in Case 1, $\widehat{\alpha}(\pi)$ is consistent, and*

$$\sqrt{nh}\left(\widehat{\alpha}(\pi) - \alpha_\pi(\pi) - B_{1n}\right) \xrightarrow{d} N\left(0, \frac{\sigma_1^2(\pi) + \sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1\right),$$

where $B_{1n} = o_p(1)$ is defined in (17), and Γ_+ and Ω_+ are defined in (4).

When the measurement error shrinks to zero, we do not require the treatment effect to be constant to identify it. The asymptotic variance is the same as the Oracle Case (3) and is similar to that in Theorem 1 except that now x is replaced by x^*. This is understandable: with a shrinking measurement error, x converges to x^* in some sense.

6.2 The Forcing Variable is x^*

As in Section 4, when d_π^* can be observed, we consider $\widetilde{\alpha}(\pi)$; otherwise, we use $\widehat{\alpha}(\pi)$.

Theorem 4 *Under Assumptions B, EY, F′, K and M, in Case 2, $\widetilde{\alpha}(\pi)$ is consistent but its asymptotic variance depends on the relative magnitude of $n^{-\delta}$ and h. If $\frac{n^{-\delta}}{h} \to 0$, then*

$$\sqrt{nh}\left(\widetilde{\alpha}(\pi) - \alpha_\pi(\pi) - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{\sigma_1^2(\pi) + \sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1\right);$$

if $\frac{n^{-\delta}}{h} \to \infty$, then

$$\sqrt{nh}\left(\widetilde{\alpha}(\pi) - \alpha_\pi(\pi) - B_{2n}\right) \xrightarrow{d} N\left(0, \frac{1}{f_{x^*}(\pi)}\left(\frac{\sigma_1^2(\pi)}{F_\epsilon(0)} + \frac{\sigma_0^2(\pi)}{1 - F_\epsilon(0)}\right) e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1\right);$$

if $\frac{n^{-\delta}}{h} \to C \in (0, \infty)$, then

$$\sqrt{nh}\left(\widetilde{\alpha}(\pi) - \alpha_\pi(\pi) - B_{2n}\right) \xrightarrow{d} N\left(0, V(C)\right),$$

where $B_{2n} = o_p(1)$ is defined in (18), and

$$\begin{aligned} V(C) \quad = \quad & \frac{\sigma_1^2(\pi)}{f_{x^*}(\pi)} e_1' \left[\overline{\Gamma}_+(C) + \Gamma_+ F_\epsilon(0)\right]^{-1} \left[\overline{\Omega}_+(C) + \Omega_+ F_\epsilon(0)\right] \left[\overline{\Gamma}_+(C) + \Gamma_+ F_\epsilon(0)\right]^{-1} e_1 \\ & + \frac{\sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1' \left[\overline{\Gamma}_-(C) + \Gamma_- \left(1 - F_\epsilon(0)\right)\right]^{-1} \left[\overline{\Omega}_-(C) + \Omega_- \left(1 - F_\epsilon(0)\right)\right] \left[\overline{\Gamma}_-(C) + \Gamma_- \left(1 - F_\epsilon(0)\right)\right]^{-1} e_1 \end{aligned}$$

with

$$\overline{\Gamma}_{+}(C) = \begin{pmatrix} \overline{\gamma}_0^{+}(C) & \cdots & \overline{\gamma}_p^{+}(C) \\ \vdots & \ddots & \vdots \\ \overline{\gamma}_p^{+}(C) & \cdots & \overline{\gamma}_{2p}^{+}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\gamma}_j^{+}(C) = \int_0^{\infty} \left[\int_{C\epsilon}^{M} k(v) v^j dv \right] f_\epsilon(\epsilon) d\epsilon,$$

$$\overline{\Omega}_{+}(C) = \begin{pmatrix} \overline{\omega}_0^{+}(C) & \cdots & \overline{\omega}_p^{+}(C) \\ \vdots & \ddots & \vdots \\ \overline{\omega}_p^{+}(C) & \cdots & \overline{\omega}_{2p}^{+}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\omega}_j^{+}(C) = \int_0^{\infty} \left[\int_{C\epsilon}^{M} k^2(v) v^j dv \right] f_\epsilon(\epsilon) d\epsilon,$$

$$\overline{\Gamma}_{-}(C) = \begin{pmatrix} \overline{\gamma}_0^{-}(C) & \cdots & \overline{\gamma}_p^{-}(C) \\ \vdots & \ddots & \vdots \\ \overline{\gamma}_p^{-}(C) & \cdots & \overline{\gamma}_{2p}^{-}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\gamma}_j^{-}(C) = \int_{-\infty}^{0} \left[\int_{-M}^{C\epsilon} k(v) v^j dv \right] f_\epsilon(\epsilon) d\epsilon,$$

$$\overline{\Omega}_{-}(C) = \begin{pmatrix} \overline{\omega}_0^{-}(C) & \cdots & \overline{\omega}_p^{-}(C) \\ \vdots & \ddots & \vdots \\ \overline{\omega}_p^{-}(C) & \cdots & \overline{\omega}_{2p}^{-}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\omega}_j^{-}(C) = \int_{-\infty}^{0} \left[\int_{-M}^{C\epsilon} k^2(v) v^l dv \right] f_\epsilon(\epsilon) d\epsilon,$$

$$\Gamma_{-} = \begin{pmatrix} \gamma_0^{-} & \cdots & \gamma_p^{-} \\ \vdots & \ddots & \vdots \\ \gamma_p^{-} & \cdots & \gamma_{2p}^{-} \end{pmatrix}_{(p+1)\times(p+1)}, \gamma_j^{-} = (-1)^j \gamma_j^{+},$$

$$\Omega_{-} = \begin{pmatrix} \omega_0^{-} & \cdots & \omega_p^{-} \\ \vdots & \ddots & \vdots \\ \omega_p^{-} & \cdots & \omega_{2p}^{-} \end{pmatrix}_{(p+1)\times(p+1)}, \omega_j^{-} = (-1)^j \omega_j^{+}.$$

We now compare the data usage of $\widehat{\alpha}(\pi)$ in Case 1 and $\widetilde{\alpha}(\pi)$ when $\frac{n^{-\delta}}{h} \to C \in (0,\infty)$. $C = 0$ and ∞ can be treated as extreme cases. Consider only $\widehat{m}_{+}(\pi)$ in $\widehat{\alpha}(\pi)$ and $\widetilde{m}_{+}(\pi)$ in $\widetilde{\alpha}(\pi)$ since similar analysis applies to $\widehat{m}_{-}(\pi)$ and $\widetilde{m}_{-}(\pi)$. Pick $x \in [\pi, \pi + h)$. In $\widetilde{m}_{+}(\pi)$, the x^* corresponding to x must satisfy that $x^* + u = x$, and $x^* \geq \pi$. So u can be positive but restricted such that $x - u \geq \pi$, which corresponds to the terms $\overline{\Gamma}_{+}(C)$ and $\overline{\Omega}_{+}(C)$ in $V(C)$, or u is negative and unrestricted, which corresponds to the terms $\Gamma_{+} F_\epsilon(0)$ and $\Omega_{+} F_\epsilon(0)$. However, in $\widehat{m}_{+}(\pi)$, the measurement error u in x is not restricted as long as $x^* + u = x$. As a result, the density of ϵ is truncated in $V(C)$, while in the asymptotic variance of $\widehat{\alpha}(\pi)$, it is integrated out and disappears. The discussion above is intuitively illustrated in the upper two panels of Figure 5.

It is easy to check that $V(C)$ converges to $\frac{\sigma_1^2(\pi)+\sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1' \Gamma_{+}^{-1} \Omega_{+} \Gamma_{+}^{-1} e_1 \equiv V(0)$ when $C \to 0$ and converges to $V(\infty) \equiv \frac{1}{f_{x^*}(\pi)} \left(\frac{\sigma_1^2(\pi)}{F_\epsilon(0)} + \frac{\sigma_0^2(\pi)}{1-F_\epsilon(0)} \right) e_1' \Gamma_{+}^{-1} \Omega_{+} \Gamma_{+}^{-1} e_1$ when $C \to \infty$. The former is much expected since when the measurement error is extremely small, the same part of data are used in estimating $\widetilde{\alpha}(\pi)$ and $\widehat{\alpha}(\pi)$. When $C \to \infty$, only the part of data with $\epsilon < 0$ are used in $\widetilde{m}_{+}(\pi)$, since the data with $\epsilon > 0$ cannot be covered by the kernel with bandwidth h in $\widetilde{m}_{+}(\pi)$. This is why $F_\epsilon(0)$ appears associated with $\sigma_1^2(\pi)$. Intuitively, the asymptotic variance of $\widetilde{\alpha}(\pi)$ should be larger than that of $\widehat{\alpha}(\pi)$ since less data are used in $\widetilde{\alpha}(\pi)$. Define the relative efficiency as

$$RE(C) = \frac{V(C)}{V(0)}.$$

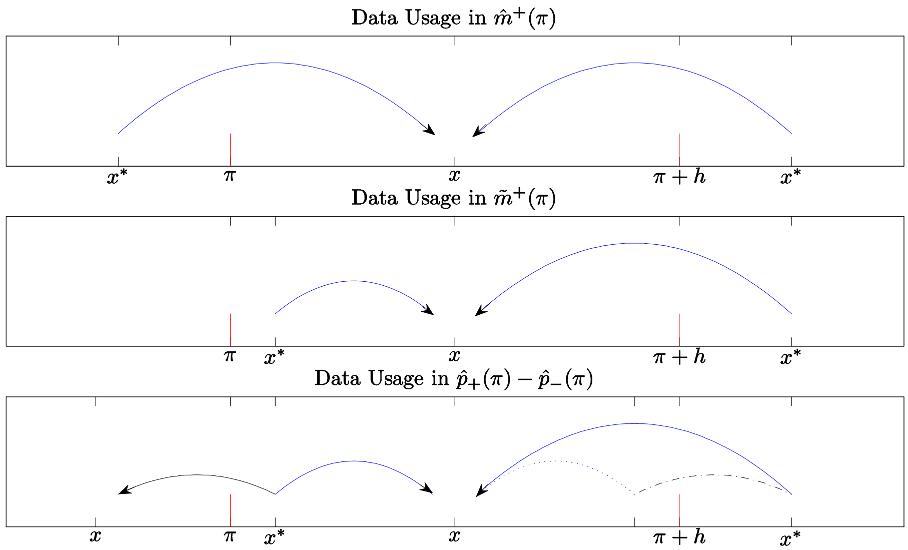

Figure 5: Data Usage in $\widehat{m}_+(\pi)$, $\widetilde{m}_+(\pi)$ and $\widehat{p}_+(\pi)-\widehat{p}_-(\pi)$

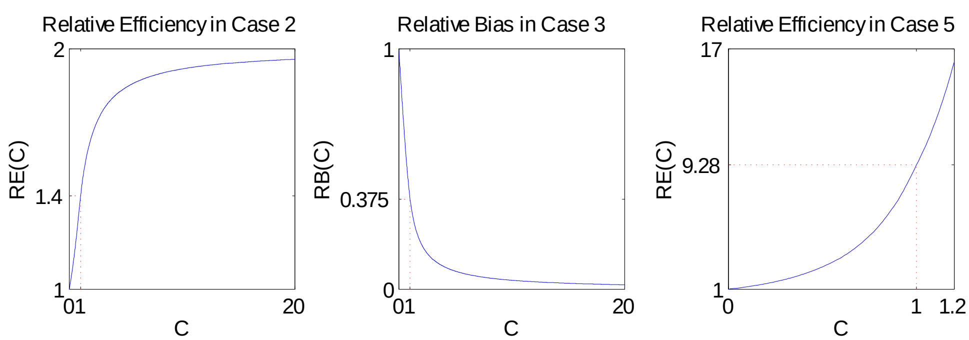

Figure 6: Relative Efficiency in Case 2 and 5 and Relative Bias in Case 3

To aid intuition, suppose ϵ is symmetrically distributed, then $V(C)$ reduces to

$$\frac{\sigma_1^2(\pi)+\sigma_0^2(\pi)}{f_{x^*}(\pi)}e_1'\left[\overline{\Gamma}_+(C)+\Gamma_+/2\right]^{-1}\left[\overline{\Omega}_+(C)+\Omega_+/2\right]\left[\overline{\Gamma}_+(C)+\Gamma_+/2\right]^{-1}e_1,$$

and $\frac{\sigma_1^2(\pi)}{F_\epsilon(0)}+\frac{\sigma_0^2(\pi)}{1-F_\epsilon(0)}$ in $V(\infty)$ reduces to $2\left(\sigma_0^2(\pi)+\sigma_1^2(\pi)\right)$. Suppose further that $p=0$, ϵ follows a uniform distribution on $[-1,1]$ and $k(v)$ is the Epanechinikov kernel, then $RE(C)$ as a function of C is shown in the left panel of Figure 6.[8] So when x^* is the forcing variable and d_π^* can be observed, the treatment effect can be identified but some efficiency is lost unless $\frac{n^{-\delta}}{h}\to 0$.

Now, we state the consistency result of $\widehat{\alpha}(\pi)$ in Case 3. Unlike the previous two cases, there are double contaminations in the estimation. The first is from mismeasuring $m(x^*)$ as $m(x)$ as in Case 1 and 2, and the second is from mixing the treated group with the controlled group.

[8]For $p=1$, $RE(C)$ is not monotone as a function of C, but always greater than 1. For some range of C values, $RE(C)$ can even be larger than 10000.

Theorem 5 *Under Assumptions B, EY, F', K and M, in Case 3, the consistency of* $\widehat{\alpha}(\pi)$ *depends on the relative magnitude of* $n^{-\delta}$ *and* h*:*

$$\widehat{\alpha}(\pi) \xrightarrow{p} \begin{cases} \alpha_\pi(\pi), & \text{if } \frac{n^{-\delta}}{h} \to 0, \\ \alpha_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^{-}(C)\right], & \text{if } \frac{n^{-\delta}}{h} \to C \in (0,\infty), \\ 0, & \text{if } \frac{n^{-\delta}}{h} \to \infty, \end{cases}$$

where

$$\overline{\mu}_{r,q}^{+}(C) = \begin{pmatrix} \overline{\gamma}_r^{+}(C) \\ \vdots \\ \overline{\gamma}_q^{+}(C) \end{pmatrix} \text{ and } \overline{\mu}_{r,q}^{-}(C) = \begin{pmatrix} \overline{\gamma}_r^{-}(C) \\ \vdots \\ \overline{\gamma}_q^{-}(C) \end{pmatrix}$$

with $r \leq q$ *being nonnegative integers. When* $\frac{n^{-\delta}}{h} \to 0$,

$$\sqrt{nh}\left(\widehat{\alpha}(\pi) - \alpha_\pi(\pi) - B_{3n}\right) \xrightarrow{d} N\left(0, \frac{\sigma_1^2(\pi) + \sigma_0^2(\pi)}{f_{x^*}(\pi)} e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1\right),$$

where $B_{3n} = o_p(1)$ *is defined in (19).*

From Theorem 5, when d_π^* cannot be observed, $\widehat{\alpha}(\pi)$ is not consistent unless $\frac{n^{-\delta}}{h} \to 0$. $e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^{-}(C)$ converges to 1 when $C \to 0$ and converges to zero when $C \to \infty$, so the bias gets larger when C gets larger. When $\frac{n^{-\delta}}{h} \to \infty$ (that is, the measurement error is large relative to the bandwidth), $\widehat{\alpha}(\pi)$ always converges to zero as in the case with fixed measurement error. The middle panel of Figure 6 shows the bias as a function of C in the example following Theorem 4. Here, we define the relative bias as

$$RB(C) = e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^{-}(C),$$

which is equal to $2e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C)$ when ϵ is symmetrically distributed.

It is interesting to note that the limit of $\widehat{\alpha}(\pi)$ is proportional to $\alpha_\pi(\pi)$ and contrast to Case 2, it is not related to $m_\pi(\cdot)$. This is easy to see from (9): $E[y|x] = E[m_\pi^*(x^*)|x] + \alpha_\pi(\pi)p(x)$. Since $E[m_\pi^*(x^*)|x]$ is continuous, $\widehat{\alpha}(\pi)$, as an estimator of $E[y|x = \pi+] - E[y|x = \pi-]$, must converge to $\alpha_\pi(\pi)(p_+(\pi) - p_-(\pi))$, a proportion of $\alpha_\pi(\pi)$. $e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^{-}(C)$ is exactly the limit of $\widehat{p}_+(\pi) - \widehat{p}_-(\pi)$ when the measurement error is shrinking. The third panel of Figure 5 shows the data usage in $\widehat{p}_+(\pi) - \widehat{p}_-(\pi)$. $\widehat{p}_+(\pi)$ uses data associated with the solid blue line, and $\widehat{p}_+(\pi)$ uses data associated with the solid black line. Part of the data used in $\widehat{p}_+(\pi)$ and $\widehat{p}_+(\pi)$ offset each other. The remaining data are associated with the (left) solid blue line, corresponding to $e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^{+}(C)$, and the dotted blue line, corresponding to $e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^{-}(C)$.

6.3 Further Analysis of Biases

When the measurement error shrinks to zero, $\alpha_\pi(\pi)$ can be identified in many cases as stated in Theorem 3, 4 and 5. But we did not study the properties of B_{1n}, B_{2n} and B_{3n} explicitly in these theorems except stating that they are $o_p(1)$. This is mainly because the exact bias properties depend on the converging process of $\frac{n^{-\delta}}{h}$ and also the smoothness of $f_{x^*}(x^*)$, $m_\pi(x^*)$ and $\alpha_\pi(x^*)$ in a complicated way in the general case. To provide further insights on the biases, we will concentrate on some interesting cases below. First, we use the LLS as in Hahn et al. (2001). Second, we assume $\frac{n^{-\delta}}{h} = C \in (0,\infty)$; that is, the measurement error is comparable with the bandwidth in an exact ratio. We can treat a small measurement error as C close to zero and a large measurement error as C close to ∞, and neglect the converging process of $\frac{n^{-\delta}}{h}$. We put further restrictions on the distribution of ϵ and the smoothness of $f_{x^*}(x^*)$, $m_\pi(x^*)$ and $\alpha_\pi(x^*)$ as follows.

Assumption F$''$: $u = n^{-\delta}\epsilon$, $\delta > 0$. ϵ has a compact support, and $f_\epsilon(\epsilon)$ is continuous on its support. $f_{x^*}(x^*)$ is continuously differentiable on $\mathcal{N}$, where $\mathcal{N}$ is a neighborhood of π, and $f_{x^*}(\pi)$ is bounded away from 0.

Assumption CY: $m_\pi(x^*)$ and $\alpha_\pi(x^*)$ are twice continuously differentiable on $\mathcal{N}$.

The compact support of ϵ in Assumption F$''$ is used to simplify proof. It can be treated as an approximation of the real measurement error. In the pretest-posttest example or the U.S. house election example, this assumption is reasonable since the recorded score or the democratic vote share margin of victory should be in a bounded range. Assumption F$''$ and CY impose more smoothness on $f_{x^*}(x^*)$, $m_\pi(x^*)$ and $\alpha_\pi(x^*)$ than Assumption F$'$ and M but only in a neighborhood of π, which is because we assume ϵ has a compact support.

Theorem 6 *Suppose Assumptions B, CY, F$''$ and K hold, and* $\frac{n^{-\delta}}{h} = C \in (0, \infty)$. *In Case 1, if the LLS is used as* $\widehat{\alpha}(\pi)$,

$$B_{1n} = \left\{ -n^{-\delta}\alpha'_\pi(\pi) E[\epsilon] + n^{-2\delta}\left[\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\alpha'_\pi(\pi) Var(\epsilon) + \frac{\alpha''_\pi(\pi)}{2} E\left[\epsilon^2\right]\right] + h^2 \frac{\alpha''_\pi(\pi)}{2} e_1'\Gamma_+^{-1}\mu_{2,3}^+ \right\}(1 + o_p(1)).$$

In Case 2, if the LLS is used as $\widetilde{\alpha}(\pi)$,

$$\begin{aligned} B_{2n} &= -n^{-\delta}\Big\{ \left(m'_\pi(\pi) + \alpha'_\pi(\pi)\right) e_1' \left[\overline{\Gamma}_+(C) + \Gamma_+ F_\epsilon(0)\right]^{-1} \left[\overline{\overline{\mu}}_{0,1}^{+}(C) + \mu_{0,1}^+ E[\epsilon|\epsilon < 0] F_\epsilon(0)\right] \\ &\quad - m'_\pi(\pi) e_1' \left[\overline{\Gamma}_-(C) + \Gamma_-(1 - F_\epsilon(0))\right]^{-1} \left[\overline{\overline{\mu}}_{0,1}^{-}(C) + \mu_{0,1}^- E[\epsilon|\epsilon > 0](1 - F_\epsilon(0))\right]\Big\}(1 + o_p(1)) \end{aligned}$$

where

$$\begin{aligned} \overline{\overline{\mu}}_{r,q}^{+}(C) &= \begin{pmatrix} \overline{\overline{\gamma}}_r^{+}(C) \\ \vdots \\ \overline{\overline{\gamma}}_q^{+}(C) \end{pmatrix}, \ \overline{\overline{\gamma}}_j^{+}(C) = \int_0^\infty \left[\int_{C\epsilon}^M k(v)v^j dv\right] \epsilon f_\epsilon(\epsilon) d\epsilon, \\ \overline{\overline{\mu}}_{r,q}^{-}(C) &= \begin{pmatrix} \overline{\overline{\gamma}}_r^{-}(C) \\ \vdots \\ \overline{\overline{\gamma}}_q^{-}(C) \end{pmatrix}, \ \overline{\overline{\gamma}}_j^{-}(C) = \int_{-\infty}^0 \left[\int_{-M}^{C\epsilon} k(v)v^j dv\right] \epsilon f_\epsilon(\epsilon) d\epsilon, \end{aligned}$$

with $r \le q$ *being nonnegative integers.*

A key difference between B_{1n} and B_{2n} is that B_{1n} is only related to $\alpha_\pi(\cdot)$, while B_{2n} is also related to $m_\pi(\cdot)$. This is not surprising by noting the comment after Theorem 2: $\overline{\alpha}_\pi(\pi)$ is only affected by $\alpha_\pi(\cdot)$, while $\Delta\overline{m}_\pi(\pi)$ is also affected by $m_\pi(\cdot)$. This fact makes the behaviors of B_{1n} and B_{2n} very different. First check B_{1n}. $h^2\frac{\alpha''_\pi(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+$ is the bias of the LLS without measurement error. The other two terms are from the measurement error. So the measurement error indeed contributes to the bias of $\widehat{\alpha}(\pi)$ although does not affect its variance. Since B_{2n} is related to C in a complicated form, it is interesting to check some special cases to sharpen our understanding. Let $C = \infty$, then

$$B_{2n} = \left\{-n^{-\delta}\left[\left(m'_\pi(\pi) + \alpha'_\pi(\pi)\right) E[\epsilon|\epsilon < 0] - m'_\pi(\pi) E[\epsilon|\epsilon > 0]\right]\right\}(1 + o_p(1)); \tag{13}$$

that is, all terms associated with C diminish. Now, it is quite clear that B_{2n} is larger than B_{1n} in many interesting cases. For example, it is common to assume $E[\epsilon] = 0$. In this case, $B_{1n} = O_p\left(n^{-2\delta} + h^2\right) = O_p\left(n^{-2\delta}\right)$, while $B_{2n} = O_p\left(n^{-\delta}\right)$. For another example, suppose $\alpha_\pi(\cdot)$ is constant, then $B_{1n} = o_p\left(n^{-2\delta} + h^2\right) = o_p\left(n^{-2\delta}\right)$, but B_{2n} is still $O_p(n^{-\delta})$ unless $m'_\pi(\pi) = 0$. In other words, B_{2n} is always in the magnitude of the measurement error, while B_{1n} can be much reduced in many cases encountered in practice.

We briefly explain why this can happen by following the discussion after Theorem 4. For $\widehat{m}_+(\pi)$, the measurement error u in $x \in [\pi, \pi + h)$ is not restricted, so the moments of ϵ appearing in B_{1n} are always integrations on the whole support of ϵ. In contrast, for $\widetilde{m}_+(\pi)$, the positive u is restricted, corresponding to the terms $\overline{\Gamma}_+(C)$ and $\overline{\overline{\mu}}^+_{0,1}(C)$ in B_{2n}, and the negative u is unrestricted, corresponding to the terms $\Gamma_+ F_\epsilon(0)$ and $\mu^+_{0,1}E[\epsilon|\epsilon < 0]F_\epsilon(0)$. Even if we assume $E[\epsilon] = 0$, it is hard to believe that the truncated expectation of ϵ is zero. Reconsider the simplified form of B_{2n}, (13). Suppose ϵ is symmetrically distributed (and $\alpha_\pi(\cdot)$ is constant), then (13) reduces to

$$B_{2n} = \left\{2n^{-\delta} m'_\pi(\pi) E[\epsilon|\epsilon > 0]\right\}(1 + o_p(1)),$$

which should be $O_p(n^{-\delta})$ in general. Combining with Theorem 5, it is reasonable to claim that the bias property of Case 1 is better than Case 2, which is in turn better than Case 3.

In Case 1, we can assess the relative magnitude of the bias from the measurement error to the bias without measurement error. For example, suppose $E[\epsilon] = 0$, then this relative magnitude equals

$$RM = \frac{Var(u)}{h^2} \frac{\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\alpha'_\pi(\pi) + \frac{\alpha''_\pi(\pi)}{2}}{\frac{\alpha''_\pi(\pi)}{2} e'_1 \Gamma_+^{-1} \mu^+_{2,3}}.$$

$f'_{x^*}(\pi)$, $f_{x^*}(\pi)$, $\alpha'_\pi(\pi)$, and $\alpha''_\pi(\pi)$ can be consistently estimated using the contaminated data if the measurement error shrinks to zero just as in consistently estimating $\widehat{\alpha}(\pi)$. Then by varying $\frac{Var(u)}{h^2}$, we can check the sensitivity of $\widehat{\alpha}(\pi)$ to the measurement error. Anyway, because estimating derivatives of $f_{x^*}(\cdot)$ and $\alpha_\pi(\cdot)$ will introduce extra biases, we do not pursue this point in this paper; see Battistin and Chesher (2009) for parallel analysis in the average treatment framework. Also note that this relative magnitude in Case 2 cannot be assessed due to the reason discussed in the last paragraph.

	Shrinking Measurement Error	Fixed Measurement Error
Case 1	Identified With a Small Extra Bias, and Without Efficiency Loss	Identified in Special Cases
Case 2	Identified With a Large Bias, and With Efficiency Loss Unless $\frac{n^{-\delta}}{h} \to 0$	Unidentified
Case 3	Unidentified Unless $\frac{n^{-\delta}}{h} \to 0$	Unidentified

Table 1: Identification for the Treatment Effect in the Sharp Design

The identification results for the treatment effect in Section 3, 4 and 5 are summarized in Table 1. The picture is clear: the identification problem for the treatment effect is harder when x^* is the forcing variable than when x is the forcing variable; when x^* is the forcing variable, the problem is harder in the case where d^*_π cannot be observed than in the case where d^*_π can be observed.

7 Fuzzy Design

In the fuzzy design, we use $\widehat{\alpha}_f(\pi)$ to estimate the treatment effect $\alpha_\pi(\pi)$ in both Case 4 and Case 5. Note that Case 4 includes Case 1 as a special case and Case 5 includes Case 2 as a special case, so the identification results below also apply to Case 1 and Case 2 when $\widehat{\alpha}_f(\pi)$ (instead of $\widehat{\alpha}(\pi)$ and $\widetilde{\alpha}(\pi)$) is used to identify $\alpha_\pi(\pi)$.

The following assumptions are imposed in the fuzzy design in addition to the assumptions in the sharp design. For ease of exposition, we define $\eta = \eta_1$ when $x \geq \pi$ in Case 4 or $x^* \geq \pi$ in Case 5, and $\eta = \eta_0$ when $x < \pi$ in Case 4 or $x^* < \pi$ in Case 5. Also, extend the conditional distribution of η_1 and η_0 continuously to all x in Case 4 and for all x^* in Case 5.

Assumption P: $p_\pi(x^*)$ and $\beta_\pi(x^*)$ are continuous on $\mathcal{X}^*$. $\beta_\pi(\pi) \neq 0$.

Assumption ED: $E\left[\eta_t\left(\varepsilon_1 - \varepsilon_0\right)|x^*\right] = 0$, $E[\eta_t|x^*] = 0$. $E[\eta_t\varepsilon_d|x^*]$, $E[\eta_t\varepsilon_d^2\big|\, x^*]$, $E[\eta_t\varepsilon_0\varepsilon_1|\, x^*]$, $E[\eta_t^2\varepsilon_d|x^*]$ and $E\left[\eta_t^2\left(\varepsilon_1 - \varepsilon_0\right)^2\Big|\, x^*\right]$ are continuous and bounded on $\mathcal{X}^*$, $d = 0, 1$, $t = 0, 1$. $\beta_\pi(\pi) \neq 0$.

Since D and η_t are bounded, $p_\pi(x^*)$, $\beta_\pi(x^*)$, $p_\pi(x)$, $\beta_\pi(x)$, $E\left[\eta_t^2|x\right]$ and $E\left[\eta_t^2|x^*\right]$ are automatically bounded, so we do not include such boundedness assumptions in Assumption P and ED as in Assumption M and EY. $E\left[\eta_t\left(\varepsilon_1 - \varepsilon_0\right)|x^*\right] = 0$ is the local unconfoundedness condition in Theorem 2 of Hahn et al. (2001). In Case 4, $E\left[\eta_0^2|x\right] = p_\pi(x)\left(1 - p_\pi(x)\right)$ and $E\left[\eta_1^2|x\right] = \left(p_\pi(x) + \beta_\pi(x)\right)\left(1 - p_\pi(x) - \beta_\pi(x)\right)$, so $E\left[\eta_0^2|x\right] \neq E\left[\eta_1^2|x\right]$ in general. Analogously, we expect the cross moments of η_t and ε_d in Assumption ED are not equal for $t = 0$ and $t = 1$.

Theorem 7 *Suppose the measurement error u is fixed, and Assumptions B, ED, EY, F, K, M, and P hold.*

(i) *In Case 4,*

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \overline{\alpha}_\pi(\pi) - B_n^f\right) \xrightarrow{d} N\left(0, \Sigma_f\right),$$

where $B_n^f = o_p(1)$ is defined in (20), and

$$\begin{aligned}\Sigma_f &= \frac{e_1'\Gamma_+^{-1}\Omega_+^{-1}\Gamma_+^{-1}e_1}{f_x(\pi)\,\beta_\pi^2(\pi)}\left\{\left[E\left[R_0^2|x=\pi\right] + E\left[R_1^2|x=\pi\right]\right]\right.\\ &\quad -2\overline{\alpha}_\pi(\pi)\left[E\left[R_0\eta_0|x=\pi\right] + E\left[R_1\eta_1|x=\pi\right]\right]\\ &\quad \left.+\overline{\alpha}_\pi(\pi)^2\left[E\left[\eta_0^2|x=\pi\right] + E\left[\eta_1^2|x=\pi\right]\right]\right\}.\end{aligned}$$

with

$$R_t = \eta_t\left(\varepsilon_1 - \varepsilon_0\right) + \alpha_\pi(x-u)\eta_t + \left[p_\pi(x) + \beta_\pi(x)t\right]\varepsilon_1 + \left[1 - p_\pi(x) - \beta_\pi(x)t\right]\varepsilon_0,\ t = 0, 1.$$

(ii) *In Case 5, $\widehat{\alpha}_f(\pi)$ is not a consistent estimator of $\alpha_\pi(\pi)$.*

Suppose the measurement error u is shrinking, and Assumptions B, ED, EY, F′, K, M, and P hold.

(iii) *In Case 4,*

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \overline{\alpha}_\pi(\pi) - B_{4n}\right) \xrightarrow{d} N\left(0, \Sigma_4\right),$$

where $B_{4n} = o_p(1)$ is defined in (21), and

$$\begin{aligned}\Sigma_4 &= \frac{e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1}{f_{x^*}(\pi)\,\beta_\pi^2(\pi)}\left\{\left[E\left[R_0^{*2}|x^*=\pi\right] + E\left[R_1^{*2}|x^*=\pi\right]\right]\right.\\ &\quad -2\alpha_\pi(\pi)\left[E\left[R_0^*\eta_0|x^*=\pi\right] + E\left[R_1^*\eta_1|x^*=\pi\right]\right]\\ &\quad \left.+\alpha_\pi(\pi)^2\left[E\left[\eta_0^2|x^*=\pi\right] + E\left[\eta_1^2|x^*=\pi\right]\right]\right\}.\end{aligned}$$

with

$$R_t^* = \eta_t\left(\varepsilon_1 - \varepsilon_0\right) + \alpha_\pi(x^*)\eta_t + \left[p_\pi(x^*) + \beta_\pi(x^*)t\right]\varepsilon_1 + \left[1 - p_\pi(x^*) - \beta_\pi(x^*)t\right]\varepsilon_0,\ t = 0, 1.$$

(iv) *In Case 5, when* $\frac{n^{-\delta}}{h} \to C \in [0, \infty)$,

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \alpha_\pi(\pi) - B_{5n}\right) \xrightarrow{d} N\left(0, \Sigma_5\left(C\right)\right),$$

where $B_{5n} = o_p(1)$ *is defined in (22), and*

$$\begin{aligned}\Sigma_5\left(C\right) &= \frac{1}{f_{x^*}\left(\pi\right)\beta_\pi^2(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right]^2} \\ &\quad \left\{E\left[R_0^{*2}|x^* = \pi\right]\left(\Sigma_{+-}(C) + \Sigma_{--}(C)\right) + E\left[R_1^{*2}|x^* = \pi\right]\left(\Sigma_{++}(C) + \Sigma_{-+}(C)\right)\right. \\ &\quad -2\alpha_\pi(\pi)\left[E\left[R_0^*\eta_0|x^* = \pi\right]\left(\Sigma_{+-}(C) + \Sigma_{--}(C)\right) + E\left[R_1^*\eta_1|x^* = \pi\right]\left(\Sigma_{++}(C) + \Sigma_{-+}(C)\right)\right] \\ &\quad \left. +\alpha_\pi(\pi)^2\left[E\left[\eta_0^2|x^* = \pi\right] + E\left[\eta_1^2|x^* = \pi\right]\right]e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1\right\}\end{aligned}$$

with

$$\begin{aligned}\Sigma_{+-}(C) &= e_1'\Gamma_+^{-1}\overline{\overline{\Omega}}_{+-}(C)\Gamma_+^{-1}e_1, \Sigma_{++}(C) = e_1'\Gamma_+^{-1}\overline{\overline{\Omega}}_{++}(C)\Gamma_+^{-1}e_1, \\ \Sigma_{-+}(C) &= e_1'\Gamma_-^{-1}\overline{\overline{\Omega}}_{-+}(C)\Gamma_-^{-1}e_1, \Sigma_{--}(C) = e_1'\Gamma_-^{-1}\overline{\overline{\Omega}}_{--}(C)\Gamma_-^{-1}e_1,\end{aligned}$$

and

$$\overline{\overline{\Omega}}_{+-}(C) = \begin{pmatrix} \overline{\overline{\omega}}_0^{+-}(C) & \cdots & \overline{\overline{\omega}}_p^{+-}(C) \\ \vdots & \ddots & \vdots \\ \overline{\overline{\omega}}_p^{+-}(C) & \cdots & \overline{\overline{\omega}}_{2p}^{+-}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\overline{\omega}}_j^{+-}(C) = \int_0^\infty \int_{-C\epsilon}^0 k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon,$$

$$\overline{\overline{\Omega}}_{++}(C) = \begin{pmatrix} \overline{\overline{\omega}}_0^{++}(C) & \cdots & \overline{\overline{\omega}}_p^{++}(C) \\ \vdots & \ddots & \vdots \\ \overline{\overline{\omega}}_p^{++}(C) & \cdots & \overline{\overline{\omega}}_{2p}^{++}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \begin{aligned}\overline{\overline{\omega}}_j^{++}(C) &= \textstyle\int_0^\infty \int_0^M k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon \\ &\quad + \textstyle\int_{-\infty}^0 \int_{-C\epsilon}^M k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon,\end{aligned}$$

$$\overline{\overline{\Omega}}_{-+}(C) = \begin{pmatrix} \overline{\overline{\omega}}_0^{-+}(C) & \cdots & \overline{\overline{\omega}}_p^{-+}(C) \\ \vdots & \ddots & \vdots \\ \overline{\overline{\omega}}_p^{-+}(C) & \cdots & \overline{\overline{\omega}}_{2p}^{-+}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \overline{\overline{\omega}}_j^{-+}(C) = \int_{-\infty}^0 \int_0^{-C\epsilon} k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon,$$

$$\overline{\overline{\Omega}}_{--}(C) = \begin{pmatrix} \overline{\overline{\omega}}_0^{--}(C) & \cdots & \overline{\overline{\omega}}_p^{--}(C) \\ \vdots & \ddots & \vdots \\ \overline{\overline{\omega}}_p^{--}(C) & \cdots & \overline{\overline{\omega}}_{2p}^{--}(C) \end{pmatrix}_{(p+1)\times(p+1)}, \begin{aligned}\overline{\overline{\omega}}_j^{--}(C) &= \textstyle\int_0^\infty \int_{-M}^{-C\epsilon} k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon \\ &\quad + \textstyle\int_{-\infty}^0 \int_{-M}^0 k^2\left(v + C\epsilon\right)\left(v + C\epsilon\right)^j f_\epsilon(\epsilon)dvd\epsilon;\end{aligned}$$

when $\frac{n^{-\delta}}{h} \to \infty$, $\widehat{\alpha}_f(\pi)$ *is not a consistent estimator of* $\alpha_\pi(\pi)$.

From Theorem 3 and 5, the probability limits of $\widehat{\alpha}_f(\pi)$ are much expected except in (iv) when $\frac{n^{-\delta}}{h} \to C \in (0, \infty)$. From Theorem 5, the numerator of $\widehat{\alpha}_f(\pi)$ converges to $\beta_\pi(\pi)\alpha_\pi(\pi)\left(e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right)$, and the denominator converges to $\beta_\pi(\pi)(e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C) + e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C))$. So the biases in the numerator and denominator offset each other, and we still get a consistent estimator. When the measurement error is large, both the numerator and the denominator of $\widehat{\alpha}_f(\pi)$ converge to zero, so the limit of $\widehat{\alpha}_f(\pi)$ is undefined. Comparing with the sharp designed case, the identification power of $\widehat{\alpha}_f(\pi)$ is weaker than $\widetilde{\alpha}(\pi)$ in Case 2 but stronger than $\widehat{\alpha}(\pi)$ in Case 3.

Σ_4 is the asymptotic variance derived in Proposition 1 and Corollary 1(c) of Porter (2003). It reduces to the asymptotic variance in Theorem 3 when the design is sharp. $\Sigma_5\left(C\right)$ converges to Σ_4 when $C \to 0$.

When $f_\epsilon(\cdot)$ is symmetric, $\Sigma_{+-}(C) = \Sigma_{-+}(C)$, and $\Sigma_{-+}(C) = \Sigma_{--}(C)$, so $\Sigma_5(C)$ can be simplified as

$$
\begin{aligned}
\Sigma_5(C) &= \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4 f_{x^*}(\pi)\, \beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^+(C)\right)^2} \left\{ E\left[R_0^{*2} | x^* = \pi\right] + E\left[R_1^{*2} | x^* = \pi\right] - \right. \\
&\quad \left. -2\alpha_\pi(\pi) \left[E\left[R_0^* \eta_0 | x^* = \pi\right] + E\left[R_1^* \eta_1 | x^* = \pi\right]\right] \right\} \\
&\quad + \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1 \alpha_\pi(\pi)^2}{4 f_{x^*}(\pi)\, \beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^+(C)\right)^2} \left\{ E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right] \right\}.
\end{aligned}
$$

To compare $\Sigma_5(C)$ with Σ_4, we assume ε_0, ε_1, η_0 and η_1 are independent of each other conditional on x^*, and $\sigma_0^2(\pi) = \sigma_1^2(\pi) = \sigma^2$. Then by some tedious calculation in supplementary materials,

$$
\begin{aligned}
\Sigma_5(C) &= \frac{2\sigma^2 - \alpha_\pi^2(\pi) \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right]}{f_{x^*}(\pi)\, \beta_\pi^2(\pi)} \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4\left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^+(C)\right)^2} \\
&\quad + \frac{\alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right]}{f_{x^*}(\pi)\, \beta_\pi^2(\pi)} \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1}{4\left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^+(C)\right)^2}, \\
\Sigma_5(0) &= \Sigma_4 = \frac{2\sigma^2}{f_{x^*}(\pi)\, \beta_\pi^2(\pi)} e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1.
\end{aligned}
$$

It is obvious that Σ_4 is greater than the asymptotic variance in the sharp design by a factor $\frac{1}{\beta_\pi^2(\pi)}$; see Theorem 3. The right panel of Figure 6 shows the relative efficiency $RE(C) = \frac{\Sigma_5(C)}{\Sigma_5(0)}$ in the example following Theorem 4. Here, we further assume $p_\pi(\pi) = 0.25$ and $\beta_\pi(\pi) = 0.5$ to calculate $E\left[\eta_0^2 | x^* = \pi\right]$ and $E\left[\eta_1^2 | x^* = \pi\right]$, and $\frac{\sigma^2}{\alpha_\pi^2(\pi)}$ is set as 1. Because $e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^+(C)$ converges to 0 as $C \to \infty$, $RE(C)$ diverges to infinity very quickly.

As in Theorem 6, we study the bias properties of $\widehat{\alpha}_f(\pi)$ when the LLS is used. For this purpose, we impose further restrictions on $p_\pi(\cdot)$ and $\beta_\pi(\cdot)$.

Assumption CD: $p_\pi(x^*)$ and $\beta_\pi(x^*)$ are twice continuously differentiable on $\mathcal{N}$. $\beta_\pi(\pi) \neq 0$.

Theorem 8 *Suppose Assumptions B, CD, CY, F$''$ and K hold, $\frac{n^{-\delta}}{h} = C \in (0, \infty)$ and the LLS is used as $\widehat{\alpha}_f(\pi)$. In Case 4,*

$$
\begin{aligned}
B_{4n} &= \left\{ -n^{-\delta} \alpha_\pi'(\pi) E[\epsilon] + n^{-2\delta} \left[\frac{f_{x^*}'(\pi)}{f_{x^*}(\pi)} \alpha_\pi'(\pi) Var(\epsilon) + \frac{\alpha_\pi''(\pi)}{2} E\left[\epsilon^2\right] \right] \right. \\
&\quad \left. + h^2 \left[\frac{\alpha_\pi''(\pi)}{2} + \frac{\beta_\pi'(\pi) \alpha_\pi'(\pi)}{\beta_\pi(\pi)} \right] e_1' \Gamma_+^{-1} \mu_{2,3}^+ \right\} (1 + o_p(1)).
\end{aligned}
$$

In Case 5,

$$
B_{5n} = \alpha_\pi'(\pi) \left[h \frac{e_1' \Gamma_+^{-1} \overline{\mu}_{1,2}^+(C) + e_1' \Gamma_-^{-1} \overline{\mu}_{1,2}^-(C)}{e_1' \Gamma_+^{-1} \overline{\mu}_{0,1}^+(C) + e_1' \Gamma_-^{-1} \overline{\mu}_{0,1}^-(C)} - n^{-\delta} \frac{e_1' \Gamma_+^{-1} \overline{\overline{\mu}}_{0,1}^+(C) + e_1' \Gamma_-^{-1} \overline{\overline{\mu}}_{0,1}^-(C)}{e_1' \Gamma_+^{-1} \overline{\mu}_{0,1}^+(C) + e_1' \Gamma_-^{-1} \overline{\mu}_{0,1}^-(C)} \right] (1 + o_p(1))
$$

The h^2 term in B_{4n} is the bias of the LLS without measurement error. It includes an extra term $\frac{\beta_\pi'(\pi)\alpha_\pi'(\pi)}{\beta_\pi(\pi)}$ relative to the sharp design case where $\beta_\pi(x) = 1$. The other two terms are contributed by the measurement error and are the same as in B_{1n}, so the measurement error does not introduce extra biases in the fuzzy design if x is the forcing variable. On the contrary, in B_{5n}, $e_1' \Gamma_+^{-1} \overline{\mu}_{0,1}^+(C) + e_1' \Gamma_-^{-1} \overline{\mu}_{0,1}^-(C)$ shrinks to zero very quickly when C gets large, so B_{5n} is expected to be large in practice. Table 2 summarizes the identification results in the fuzzy design.

	Shrinking Measurement Error	Fixed Measurement Error
Case 4	Identified With a Small Extra Bias, and Without Efficiency Loss	Identified in Special Cases
Case 5	Identified Unless $\frac{n^{-\delta}}{h} \to \infty$, But With a Large Bias and a Large Variance	Unidentified

Table 2: Identification for the Treatment Effect in the Fuzzy Design

8 Monte Carlo Results

In practice, the main concern is whether the LPE is an unbiased estimator of $\alpha_\pi(\pi)$ when there is a measurement error in x^*. The goal of this simulation study is to check the bias property of the LPE in a nonlinear environment. Specifically, we consider two DGPs of y; the first one corresponds to common treatment effects, and the second one corresponds to variable treatment effects. Similar setups are used in Porter and Yu (2010).

$$\begin{aligned} DGP1 &: \ y = m_\pi\left(x^*\right) + \alpha_\pi(x^*)D + \varepsilon \text{ with } m_\pi\left(x^*\right) = x^{*2} \text{ and } \alpha_\pi(x^*) = 1, \\ DGP2 &: \ y = m_\pi\left(x^*\right) + \alpha_\pi(x^*)D + \varepsilon \text{ with } m_\pi\left(x^*\right) = x^{*2} \text{ and } \alpha_\pi(x^*) = -6x^* + 7. \end{aligned}$$

In both setups, ε's are i.i.d. sampled and follow $N(0, 0.2^2)$, and x^* is uniformly distributed on $[0,3]$. $x = x^* + u$, $u = s\epsilon$, and ϵ follows the standard normal distribution. The specifications of $\alpha_\pi(x^*)$ and the distributions of x^* and ϵ are also used in Example 2 of Section 3. $\pi = 1$, so $\alpha_\pi(\pi) = 1$ in both setups. $n = 400$, and the number of replications is set as 1000. Throughout the simulations, all estimators are based on the LLS with the Epanechinikov kernel $k(u) = \frac{3}{4}\left(1 - u^2\right)1\left(|u| \le 1\right)$.

Figure 7 and 8 show the biases as a function of s/h and s respectively for three bandwidths in the sharp design. From these two figures, three results of interest are summarized as follows. First, in Case 1, the bias is close to zero when the treatment effect is constant. Even when the treatment effect is variable, the bias is small for a large range of s/h (or s); for example, the bias is relatively small even when $s = 0.5$ which is considerably large compared with the standard deviation of ε. This confirms the magnitude of B_{1n} in Theorem 6: since $E[\epsilon] = 0$, $B_{1n} = o_p(s^2)$ in DGP1, and $B_{1n} = O_p(s^2)$ is quadratic in DGP2. Second, the bias property of Case 1 is better than that of Case 2 which is in turn better than that of Case 3. The range of s/h (or s) with small biases is much narrower in Case 2 than in Case 1. In Case 3, the bias deteriorates very quickly when s/h (or s) gets large. This confirms the consistency result in Theorem 5 and the magnitude of B_{2n} in Theorem 6. In Case 2, $B_{2n} = O_p(s)$, so the bias is roughly a linear function of s. In Case 3, the LLS is not even consistent. Because $\alpha_\pi(\pi) = 1$ and the bias converges to -1 in this case, $\widehat{\alpha}(\pi)$ converges to zero as expected. Third, comparing Figure 7 and 8, it seems that the bias in Case 1 and 2 is determined by the absolute measurement error s, while in Case 3, it is determined by the relative measurement error s/h. The former confirms the form of B_{1n} and B_{2n} in Theorem 6: they depend on $n^{-\delta}$ not $n^{-\delta}/h$. The latter confirms the limit of $\widehat{\alpha}(\pi)$ in Theorem 5: $RB(C)$ is a function of C not $n^{-\delta}$. Furthermore, adding 1 to the bias curves in Case 3 generates similar curves as in the middle panel of Figure 6. Figure 9 shows the performance of the bias approximation using the fixed measurement error framework in Section 3 and 4. In Case 1 and 2 under DGP1, both the fixed and shrinking measurement error framework provide good approximations of the real biases. But in other scenarios, the prediction using the fixed measurement error framework does not work well. For example, in case 1 under DGP2, the prediction is linear in s, but the real bias is approximately quadratic in s; in case 2 under DGP2, the prediction can even be positive, but

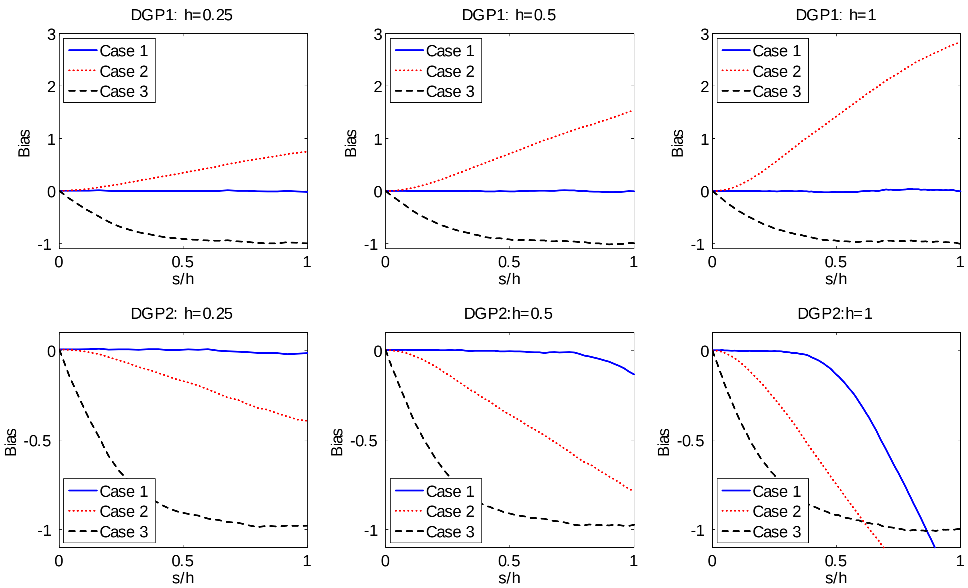

Figure 7: Bias of the LLS as a Function of s/h for Three Bandwidths in the Sharp Design

the real bias is always negative; in case 3 under both DGPs, the prediction is constant in s, but the real bias is far from constant. In summary, the approximation under the fixed measurement error framework is not useful in finite samples especially when the treatment effect is variable; on the contrary, the shrinking measurement error framework provides quite precise predictions of the biases in practice.

In the fuzzy design, when x is the forcing variable, we let

$$D = \Phi(x - 1.5)\, 1(x < 1) + \Phi\left(\frac{x - 0.3243}{0.7744}\right) 1(x \geq 1) + \eta$$

in DGP1, and

$$D = \Phi(x - 1.5)\, 1(x < 1) + \Phi(x - 0.1275)\, 1(x \geq 1) + \eta$$

in DGP2, where $\Phi(\cdot)$ is the cdf of a standard normal distribution. The former corresponds to the case where $\beta_\pi(\cdot)$ is constant ($\beta'_\pi(\pi) = 0$), while the latter corresponds to the case where $\beta_\pi(\cdot)$ is variable ($\beta'_\pi(\pi) \neq 0$). In both cases, $\beta_\pi(\pi) = 0.5$, and η is independent of ε. When x^* is the forcing variable, we just change x in the above specifications to x^*. The same bandwidth is used in the estimation of the numerator and denominator of $\widehat{\alpha}_f(\pi)$.

Figure 10 shows the bias as a function of s/h for two bandwidths in the fuzzy design. There is too much variation in the bias when $h = 0.25$, and the bias as a function of s provides similar information, so both are omitted. A few results of interest are as follows. First, from Figure 10, the bias in Case 5 is much larger than that in Case 4. This is not surprising from Theorem 8. Second, comparing Case 4 in Figure 10 with Case 1 in Figure 7, although the bias does not increase much, there seems more variation in Case 4 as predicted by (iii) of Theorem 7. Third, comparing Case 5 in Figure 10 with Case 3 in Figure 7, although $\widehat{\alpha}_f(\pi)$ is consistent in Case 5 and $\widehat{\alpha}(\pi)$ is not in Case 3, the large variance of $\widehat{\alpha}_f(\pi)$ completely ruins its performance.

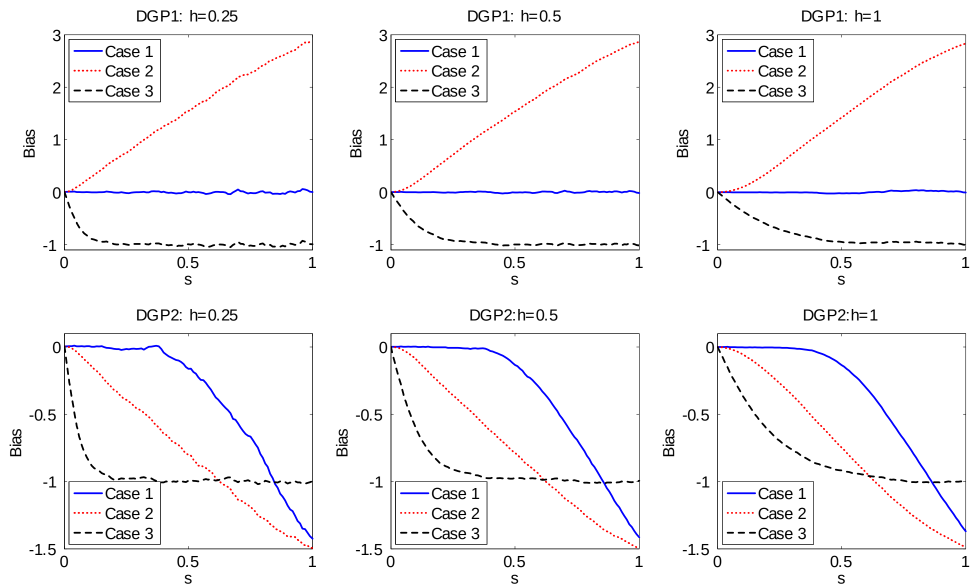

Figure 8: Bias of the LLS as a Function of s for Three Bandwidths in the Sharp Design

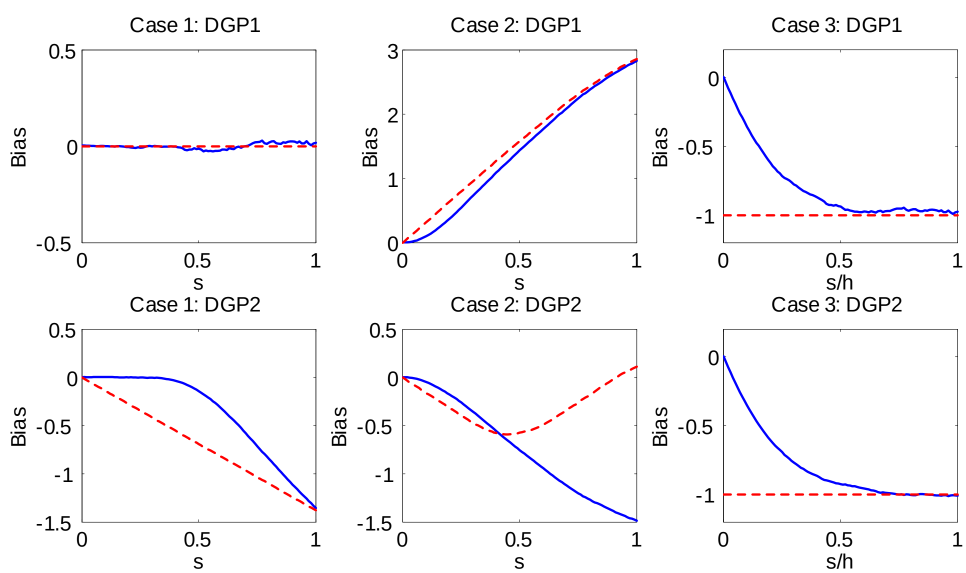

Figure 9: The Biases Predicted Under the Fixed Measurement Error Framework (Red Dashed Line): Solid Blue Lines For the Real Biases When $h = 1$ in Figure 7 and 8

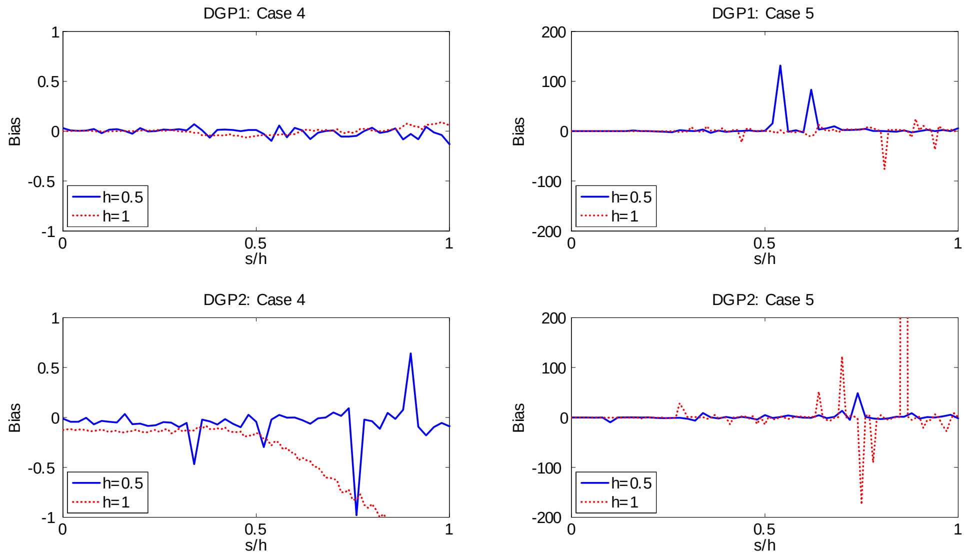

Figure 10: Bias of the LLS as a Function of s/h for Two Bandwidths in the Fuzzy Design

9 Conclusion

Measure errors are prevalent in the RDD analysis. This paper studies the identification of the treatment effect by the local polynomial estimator when the measurement error is present. The general result is that the treatment effect in the case where the treatment is based on the contaminated forcing variable is easier to identify than in the case where the treatment is based on the error free forcing variable. In practice, when the former happens, it is relatively safe to use the LPE to estimate the treatment effect. But when the latter happens, especially when the treatment status cannot be observed or in the fuzzy design, we must take caution in interpreting our estimator. As emphasized in Section 3.5.3 of Lee and Lemieux (2010), regression discontinuity should better be trearted as a "design" instead of a "method". So the analysis in this paper implies that it is better for the designer to reveal to econometrians exactly which variable the treatment is based on such that a more precise analysis on the treatment effect can be conducted.

References

Almond, D. et al., 2010, Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns, *Quarterly Journal of Economics*, 125, 591–634.

Almond, D. et al., 2011, The Role of Hospital Heterogeneity in Measuring Marginal Returns to Medical Care: A Reply to Barreca, Guldi, Lindo, and Waddell, *Quarterly Journal of Economics*, 126, 2125-2131.

Barreca, A.I. et al., 2011, Heaping-Induced Bias in Regression-Discontinuity Designs," NBER Working Paper 17408.

Barreca, A.I. et al., 2011, Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification, *Quarterly Journal of Economics*, 126, 1-7.

Battistin, E., and A. Chesher, 2009, Treatment Effect Estimation with Covariate Measurement Error, cemmap unpublished manuscript.

Bound, J., C. Brown and N. Mathiowetz, 2001, Measurement Error in Survey Data, in J.J. Heckman and E. leamer, eds., *Handbook of Econometrics*, Vol. 5, Amsterdam: North-Holland, 3705-3843.

Cappelleri, J.C. et al., 1991, Random Measurement Error Does Not Bias the Treatment Effect Estimate in the Regression-Discontinuity Design, I, The Case of No Interaction, *Evaluation Review*, 15, 395-419.

Carroll, R., D. Ruppert and L. Stefanski, 2006, *Measurement Error in Non-Linear Models*: *A Modern Perspective*, 2nd ed., London: Chapman and Hall.

Cochran, W.G. and D.B. Rubin, 1973, Controlling Bias in Observational Studies: A Review, *Sankhyā, Series A*, 35, 417-466.

Cook, T.D., 2008, "Waiting for Life to Arrive": A History of the Regression-Discontinuity Design in Psychology, Statistics and Economics, *Journal of Econometrics*, 142, 636–654.

Cook, T.D. and D.T. Campbell, 1979, *Quasi-experimentation: Design and Analysis Issues for Field Settings*, Chicago: Rand McNally.

de la Mata, D., 2011, The effect of Medicaid on Children's Health: a Regression Discontinuity Approach, Working Paper 11/16, HEDG, The University of York.

Fan, J., 1992, Design-adaptive Nonparametric Regression, *Journal of the American Statistical Association*, 87, 998-1004.

Fan, J., 1993, Local Linear Regression Smoothers and Their Minimax Efficiency, *Annals of Statistics*, 21, 196-216.

Fan, J. and Y.K. Truong, 1993, Nonparametric Regression with Errors in Variables, *Annals of Statistics*, 21, 1900-1925.

Goldberger, A.S., 2008, Selection Bias in Evaluating Treatment Effects: Some Formal Illustrations, *Advances in Econometrics*, 21, 1-31.

Hahn, J., P. Todd and W. Van der Klaauw, 2001, Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design, *Econometrica*, 69, 201–209.

Hausman, J., J. Abreveya and F. Scott-Morton, 1998, Misclassification of the Dependent Variable in a Discrete Response Setting, *Journal of Econometrics*, 87, 239-269.

Hausman, J., W. Newey, H. Ichimura and J. Powell, 1991, Identification and Estimation of Polynomial Errors-in-Variables Models, *Journal of Econometrics*, 50, 273-296.

Hausman, J., W. Newey and J. Powell, 1995, Nonlinear Errors in Variables: Estimation of Some Engel Curves, *Journal of Econometrics*, 65, 205-233.

Heckman, J., S. Urzua and E. Vytlacil, 2006, Understanding Instrumental Variables in Models with Essential Heterogeneity, *The Review of Economics and Statistics*, 389-432.

Hullegie, P. and T.J. Klein, 2010, The Effect of Private Health Insurance on Medical Care Utilization and Self-assessed Health in Germany, *Health Economics*, 19, 1048–1062.

Imbens, G.W. and T. Lemieux, 2008, Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics*, 142, 615-635.

Koch, T.G., 2010, Using RD Design to Understand Heterogeneity in Health Insurance Crowd-Out, Mimeo, University of California Santa Barbara.

Lee, D.S., 2008, Randomized Experiments from Non-random Selection in U.S. House Elections, *Journal of Econometrics*, 142, 675-697.

Lee, D.S. and T. Lemieux, 2010, Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, 48, 281-355.

Lewbel, A., 2007, Estimation of Average Treatment Effects with Misclassification, *Econometrica*, 75, 537-551.

Mahajan, A., 2006, Identification and Estimation of Regression Models with Misclassification, *Econometrica*, 74, 631-665.

Matsudaira, J.D., 2008, Mandatory Summer School and Student Achievement, *Journal of Econometrics*, 142, 829-850.

Porter, J., 2003, Estimation in the Regression Discontinuity Model, Mimeo, Department of Economics, University of Wisconsin at Madison.

Porter, J. and P. Yu, 2010, Regression Discontinuity with Unknown Discontinuity Points: Testing and Estimation, Mimeo, Department of Economics, University of Auckland.

Pei, Z., 2011, Regression Discontinuity Design with Measurement Error in the Assignment Variable, Mimeo, Department of Economics, University of Princeton.

Rubin, D.B., 1977, Assignment to Treatment Group on the Basis of a Covariate, *Journal of Educational Statistics*, 2, 1-26.

Schanzenbach, D.W., 2009, Do School Lunches Contribute to Childhood Obesity?, *Journal of Human Resources*, 44, 684–709.

Trochim, W., 1984, *Research Design for Program Evaluation: The Regression Discontinuity Approach*, Beverly Hills: Sage Publications.

Trochim, W., J.C. Cappelleri and C.S. Reichardt, 1991, Random Measurement Error Does Not Bias the Treatment Effect Estimate in the Regression-Discontinuity Design, II, When an Interaction Effect is Present, *Evaluation Review*, 15, 571-604.

Van der Klaauw, W., 2008, Regression-Discontinuity Analysis: A Survey of Recent Development in Economics, *Labour*, 22, 219-245.

Appendix: Proofs

Throughout the proofs, DCT means the dominated convergence theorem. Since the proofs are quite standard, we provide details only in the proof of Theorem 1. For the other proofs, we only give out the key steps, and omit tedious calculations.

Proof of Theorem 1. Note that

$$\sqrt{nh}\left(\widehat{\alpha}(\pi)-\overline{\alpha}_{\pi}\left(\pi\right)\right)=\sqrt{nh}\left(\widehat{m}_{+}(\pi)-\widehat{m}_{-}(\pi)-\overline{\alpha}_{\pi}\left(\pi\right)\right), \tag{14}$$

where

$$\begin{aligned}
\widehat{m}_{+}(\pi) &= e_1'\left(\frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)Z_i'\left(\pi\right)d_i(\pi)k_h\left(x_i-\pi\right)\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)d_i(\pi)k_h\left(x_i-\pi\right)y_i\right) \qquad (15)\\
&\equiv e_1'S_{n+}^{-1}(\pi)\widetilde{r}_{+}\left(y(\pi)\right),\\
\widehat{m}_{-}(\pi) &= e_1'\left(\frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)Z_i'\left(\pi\right)\left(1-d_i(\pi)\right)k_h\left(x_i-\pi\right)\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)\left(1-d_i(\pi)\right)k_h\left(x_i-\pi\right)y_i\right)\\
&\equiv e_1'S_{n-}^{-1}(\pi)\widetilde{r}_{-}\left(y(\pi)\right),
\end{aligned}$$

with

$$Z_i\left(x\right)=\left(1,\frac{x_i-x}{h},\cdots,\left(\frac{x_i-x}{h}\right)^p\right)'_{(p+1)\times 1}.$$

Since there is a measurement error in x_i^*, $y_i=m_\pi(x_i-u_i)+\alpha_\pi\left(x_i-u_i\right)+\varepsilon_{1i}$ when $d_i(\pi)=1$ and $y_i=m_\pi(x_i-u_i)+\varepsilon_{0i}$ when $d_i(\pi)=0$. Define

$$B_{1n}=e_1'S_{n+}^{-1}(\pi)\overline{r}_{+}\left(\pi\right)-e_1'S_{n-}^{-1}(\pi)\overline{r}_{-}\left(\pi\right)-\overline{\alpha}_{\pi}\left(\pi\right), \tag{16}$$

where

$$\begin{aligned}
\overline{r}_{+}\left(\pi\right) &= \frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)d_i(\pi)k_h\left(x_i-\pi\right)\left(m_\pi(x_i-u_i)+\alpha_\pi\left(x_i-u_i\right)\right),\\
\overline{r}_{-}\left(\pi\right) &= \frac{1}{n}\sum_{i=1}^{n}Z_i\left(\pi\right)\left(1-d_i(\pi)\right)k_h\left(x_i-\pi\right)m_\pi(x_i-u_i).
\end{aligned}$$

We show $B_{1n}=o_p(1)$ as follows. For $\mathfrak{p}=0,\cdots,p$,

$$\begin{aligned}
&E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}d_i(\pi)k_h\left(x_i-\pi\right)\left(m_\pi(x_i-u_i)+\alpha_\pi\left(x_i-u_i\right)\right)\right]\\
=&\int\int_0^M k(v)v^{\mathfrak{p}}\left(m_\pi(\pi+vh-u)+\alpha_\pi\left(\pi+vh-u\right)\right)f_{x^*}(\pi+vh-u)f_u(u)dvdu\\
\to&\gamma_{\mathfrak{p}}^{+}f_x(\pi)E\left[m_\pi(\pi-u)+\alpha_\pi(\pi-u)|x=\pi\right],
\end{aligned}$$

and

$$Var\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}d_i(\pi)k_h\left(x_i-\pi\right)\left(m_\pi(x_i-u_i)+\alpha_\pi\left(x_i-u_i\right)\right)\right)$$

$$\begin{aligned}
&\leq \frac{2}{nh}\int\int_0^M k^2(v)v^{2p}\left(m_\pi^2(\pi+vh-u)+\alpha_\pi^2(\pi+vh-u)\right)f_{x^*}(\pi+vh-u)f_u(u)dvdu \\
&= O\left(\frac{1}{nh}\right)=o(1).
\end{aligned}$$

by Assumption B, F, K, M, and the DCT. So $\overline{r}_+(\pi)=f_x(\pi)\left[\overline{m}_\pi(\pi)+\overline{\alpha}_\pi(\pi)\right]\mu_{0,p}^+ + o_p(1)$. Similarly, $\overline{r}_-(\pi)=f_x(\pi)\overline{m}_\pi(\pi)\mu_{0,p}^- + o_p(1)$. We then establish the probability limit for the denominator term S_{n+}. For $l=0,\cdots,2p$,

$$E\left[\frac{1}{n}\sum_{i=1}^n k_h(x_i-\pi)\,d_i(\pi)\left(\frac{x_i-\pi}{h}\right)^l\right]=\int_0^M k(v)v^l f_x(\pi+vh)\,dv\to f_x(\pi)\gamma_l^+$$

and

$$Var\left(\frac{1}{n}\sum_{i=1}^n k_h(x_i-\pi)\,d_i(\pi)\left(\frac{x_i-\pi}{h}\right)^l\right)\leq\frac{1}{nh}\int_0^M k^2(v)v^{2l}f_x(\pi+vh)\,dv\to 0$$

by Assumption B, F, K, and the DCT. So $S_{n+}=f_x(\pi)\Gamma_+ + o_p(1)$. By the continuity of matrix inversion, $S_{n+}^{-1}\xrightarrow{p} f_x^{-1}(\pi)\Gamma_+^{-1}$. Similarly, $S_{n-}^{-1}\xrightarrow{p} f_x^{-1}(\pi)\Gamma_-^{-1}$. So

$$B_{1n}\xrightarrow{p}\left(e_1'\Gamma_+^{-1}\mu_{0,p}^+\left(\overline{m}_\pi(\pi)+\overline{\alpha}_\pi(\pi)\right)-e_1'\Gamma_-^{-1}\mu_{0,p}^-\overline{m}_\pi(\pi)\right)-\overline{\alpha}_\pi(\pi)=\overline{m}_\pi(\pi)+\overline{\alpha}_\pi(\pi)-\overline{m}_\pi(\pi)-\overline{\alpha}_\pi(\pi)=0.$$

At last,

$$\begin{aligned}
&\sqrt{nh}\left(\widehat{\alpha}(\pi)-\overline{\alpha}_\pi(\pi)-B_{1n}\right)\\
=\;& e_1'S_{n+}^{-1}(\pi)\left[\frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i\right]-e_1'S_{n-}^{-1}(\pi)\left[\frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi)\left(1-d_i(\pi)\right)k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i\right]
\end{aligned}$$

We now show $\frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k\left(\frac{x_i-\pi}{h}\right)\varepsilon_{1i}\xrightarrow{d}N\left(0,f_x(\pi)\,E\left[\varepsilon_1^2|x=\pi\right]\Omega_+\right)$. The Cramer-Wold device will be applied to derive the asymptotic normality, so let λ be a nonzero, finite vector and define $U_{ni}=k\left(\frac{x_i-\pi}{h}\right)d_i(\pi)\lambda'Z_i(\pi)\,\varepsilon_{1i}/\sqrt{nh}$. $E[U_{ni}]=E[E[U_{ni}|x_i]]=0$ is proved in the main text, and for $l=0,\cdots,2p$,

$$\begin{aligned}
&\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi)\varepsilon_{1i}^2\right]\\
&= \int_0^M\int k^2(v)\,v^l\varepsilon^2\left[\int f_{x^*,\varepsilon}^1(\pi+vh-u,\varepsilon)\,f_u(u)du\right]d\varepsilon dv\\
&= \int_0^M k^2(v)\,v^l\int E\left[\varepsilon_1^2|x^*=\pi+vh-u\right]f_{x^*}(\pi+vh-u)\,f_u(u)dudv\\
&\to \int_0^M k^2(v)\,v^l dv\int E\left[\varepsilon_1^2|x^*=\pi-u\right]f_{x^*}(\pi-u)\,f_u(u)du=f_x(\pi)\,E\left[\varepsilon_1^2|x=\pi\right]\omega_l^+,
\end{aligned}$$

where the second equality is from Fubini's theorem, the convergence is from Assumption EY(b), F and the DCT, and the last equality is from a direct calculation. So

$$\sum_{i=1}^n Var(U_{ni})\to f_x(\pi)E\left[\varepsilon_1^2|x=\pi\right]\lambda'\Omega_+\lambda.$$

For $\zeta > 0$ and constants C, $C' > 0$,

$$
\begin{aligned}
& \sum_{i=1}^{n} E\left[\left|U_{ni}\right|^{2+\zeta}\right] \\
= & \sum_{i=1}^{n}\left(\frac{1}{nh}\right)^{\zeta/2}\frac{1}{nh}E\left[\left|k\left(\frac{x_i-\pi}{h}\right)\right|^{2+\zeta} d_i(\pi)\left|\lambda' Z_i(\pi)\right|^{2+\zeta}\left|\varepsilon_{1i}\right|^{2+\zeta}\right] \\
\leq & C\left(\frac{1}{nh}\right)^{\zeta/2}\frac{1}{h}E\left[\left|k\left(\frac{x-\pi}{h}\right)\right|^{2+\zeta} d_\pi E\left[\left|\varepsilon_1\right|^{2+\zeta}\middle| x\right]\sum_{\mathfrak{p}=0}^{p}\left|\lambda_{\mathfrak{p}}\left(\frac{x-\pi}{h}\right)^{\mathfrak{p}}\right|^{2+\zeta}\right] \\
\leq & C'\left(\frac{1}{nh}\right)^{\zeta/2}\sup_{x\in\mathcal{N}}E\left[\left|\varepsilon_1\right|^{2+\zeta}\middle| x\right]\int_0^M\left|k(v)\right|^{2+\zeta} f_x(\pi+vh)\sum_{\mathfrak{p}=0}^{p}\left|\lambda_{\mathfrak{p}}v^{\mathfrak{p}}\right|^{2+\zeta}dv \\
= & O\left(\left(\frac{1}{nh}\right)^{\zeta/2}\right)=o(1),
\end{aligned}
$$

where $\mathcal{N}$ is a small neighborhood of π, and the second-to-last equality is from two facts: first,

$$
\begin{aligned}
\sup_{x\in\mathcal{X}}E\left[\left|\varepsilon_1\right|^{2+\zeta}\middle| x\right] &= \sup_{x\in\mathcal{N}}\int\left|\varepsilon\right|^{2+\zeta}\frac{\int f^1_{x^*,\varepsilon}(x-u,\varepsilon)f_u(u)du}{\int f_{x^*}(x-u)f_u(u)du}d\varepsilon \\
&= \sup_{x\in\mathcal{N}}\int\int\left|\varepsilon\right|^{2+\zeta}\frac{f^1_{x^*,\varepsilon}(x-u,\varepsilon)f_u(u)}{f_x(x)}d\varepsilon du \\
&= \sup_{x\in\mathcal{N}}\int E\left[\left|\varepsilon_1\right|^{2+\zeta}\middle| x^*=x-u\right]\frac{f_{x^*}(x-u)f_u(u)}{f_x(x)}du \\
&= \sup_{x\in\mathcal{N}}\int E\left[\left|\varepsilon_1\right|^{2+\zeta}\middle| x^*=x-u\right]f(u|x)du \\
&< \infty,
\end{aligned}
$$

where the last inequality is from Assumption F and EY(c); second,

$$
\int_0^M\left|k(v)\right|^{2+\zeta} f_x(\pi+vh)\sum_{\mathfrak{p}=0}^{p}\left|\lambda_{\mathfrak{p}}v^{\mathfrak{p}}\right|^{2+\zeta}dv \to f_x(\pi)\int_0^M\left|k(v)\right|^{2+\zeta}\sum_{\mathfrak{p}=0}^{p}\left|\lambda_{\mathfrak{p}}v^{\mathfrak{p}}\right|^{2+\zeta}dv<\infty,
$$

where the convergence is from the DCT and Assumption K, F. Application of Liapunov's CLT and the Cramer-Wold device complete the argument. Similarly, we can show $\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}Z_i(\pi)(1-d_i(\pi))\,k\left(\frac{x_i-\pi}{h}\right)\varepsilon_{0i}\xrightarrow{d} N\left(0,f_x(\pi)E\left[\varepsilon_0^2|x=\pi\right]\Omega_-\right)$, where Ω_- is similarly defined as Ω_+ with ω_j^+ replaced by $\omega_j^-\equiv(-1)^j\omega_j^+$.

In summary,

$$
\sqrt{nh}\left(\widehat{\alpha}(\pi)-\overline{\alpha}_\pi(\pi)-B_{1n}\right)\xrightarrow{d} N\left(0,\frac{E\left[\varepsilon_0^2|x=\pi\right]+E\left[\varepsilon_1^2|x=\pi\right]}{f_x(\pi)}e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1\right),
$$

by noting that

$$
e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1=e_1'\Gamma_-^{-1}\Omega_-\Gamma_-^{-1}e_1,
$$

and $\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}Z_i(\pi)\,d_i(\pi)k\left(\frac{x_i-\pi}{h}\right)\varepsilon_{1i}$ and $\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}Z_i(\pi)(1-d_i(\pi))\,k\left(\frac{x_i-\pi}{h}\right)\varepsilon_{0i}$ are independent. ■

Proof of Theorem 2. Note that

$$
\widetilde{\alpha}(\pi)=\widetilde{m}_+(\pi)-\widetilde{m}_-(\pi),
$$

where

$$\begin{aligned}
\widetilde{m}_+(\pi) &= e_1'\left(\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, Z_i'(\pi)\, d_i(\pi) d_i^*(\pi) k_h(x_i-\pi)\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi) d_i^*(\pi) k_h(x_i-\pi)\, y_i\right) \\
&\equiv e_1' \underline{S}_{n+}^{-1}(\pi) \underline{\widetilde{r}}_+(y(\pi)), \\
\widetilde{m}_-(\pi) &= e_1'\left(\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, Z_i'(\pi)\,(1-d_i(\pi))\,(1-d_i^*(\pi))\, k_h(x_i-\pi)\right)^{-1} \cdot \\
&\quad \left(\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\,(1-d_i(\pi))\,(1-d_i^*(\pi))\, k_h(x_i-\pi)\, y_i\right) \equiv e_1' \underline{S}_{n-}^{-1}(\pi) \underline{\widetilde{r}}_-(y(\pi)), .
\end{aligned}$$

In this case, $y_i = m_\pi^*(x_i - u_i) + \alpha_\pi(\pi) + \varepsilon_{1i}$ in $\widetilde{m}_+(\pi)$, and $y_i = m_\pi^*(x_i - u_i) + \varepsilon_{0i}$ in $\widetilde{m}_-(\pi)$. Following the same steps as in the proof of Theorem 1, we can find the probability limit of $\widetilde{m}_+(\pi)$ and $\widetilde{m}_-(\pi)$.

For $\mathfrak{p} = 0, \cdots, p$,

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi) d_i^*(\pi) k_h(x_i-\pi)\, m_\pi^*(x_i-u_i)\right] \\
=\ & \int_0^\infty \int_{\frac{u}{h}}^{M} k(v) v^{\mathfrak{p}} m_\pi^*(\pi+vh-u) f_{x^*}(\pi+vh-u) f_u(u)\, dv du \\
& + \int_{-\infty}^{0} \int_0^M k(v) v^{\mathfrak{p}} m_\pi^*(\pi+vh-u) f_{x^*}(\pi+vh-u) f_u(u)\, dv du \\
\to\ & \gamma_{\mathfrak{p}}^+ \int_{-\infty}^{0} m_\pi^*(\pi-u) f_{x^*}(\pi-u) f_u(u)\, du \\
=\ & \gamma_{\mathfrak{p}}^+ \int_\pi^\infty m_\pi^*(x^*) f_{x^*}(x^*) f_u(\pi-x^*)\, dx^* = \gamma_{\mathfrak{p}}^+ f_x(\pi) \int_\pi^\infty m_\pi^*(x^*) f_{x^*|x}(x^*|\pi)\, dx^*,
\end{aligned}$$

and for $l = 0, \cdots, 2p$,

$$E\left[\frac{1}{n}\sum_{i=1}^{n} k_h(x_i-\pi)\, d_i(\pi) d_i^*(\pi)\left(\frac{x_i-\pi}{h}\right)^{l}\right] \to \gamma_l^+ \int_\pi^\infty f_{x^*}(x^*) f_u(\pi-x^*)\, dx^* = \gamma_l^+ f_x(\pi)\left(1 - F_{x^*|x}(\pi|\pi)\right).$$

Furthermore, their variances converge to zero by the DCT, so the limits of the means are also their probability limits. Symmetrically,

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} (1-d_i(\pi))\,(1-d_i^*(\pi))\, k_h(x_i-\pi)\, m_\pi^*(x_i-u_i) \\
\xrightarrow{p}\ & \gamma_{\mathfrak{p}}^- \int_{-\infty}^{\pi} m_\pi^*(x^*) f_{x^*}(x^*) f_u(\pi-x^*)\, dx^* = \gamma_{\mathfrak{p}}^- f_x(\pi) \int_{-\infty}^{\pi} m_\pi^*(x^*) f_{x^*|x}(x^*|\pi)\, dx^*,
\end{aligned}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} k_h(x_i-\pi)\,(1-d_i(\pi))\,(1-d_i^*(\pi))\left(\frac{x_i-\pi}{h}\right)^{l} \xrightarrow{p} \gamma_l^- \int_{-\infty}^{\pi} f_{x^*}(x^*) f_u(\pi-x^*)\, dx^* = \gamma_l^- f_x(\pi)\, F_{x^*|x}(\pi|\pi).$$

Also, it can be shown that

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)d_i^*(\pi)k_h\left(x_i-\pi\right)\varepsilon_{1i}\xrightarrow{p}0,$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}\left(1-d_i(\pi)\right)\left(1-d_i^*(\pi)\right)k_h\left(x_i-\pi\right)\varepsilon_{0i}\xrightarrow{p}0.$$

In summary,

$$\widetilde{\alpha}(\pi)\xrightarrow{p}\alpha_\pi\left(\pi\right)+e_1'\Gamma_+^{-1}\mu_{0,p}^+\overline{m}_\pi^+(\pi)-e_1'\Gamma_-^{-1}\mu_{0,p}^-\overline{m}_\pi^-(\pi)=\alpha_\pi+\Delta\overline{m}_\pi(\pi).$$

As to $\widehat{\alpha}(\pi)$,

$$y_i=d_i^*(\pi)\left(m_\pi^*(x_i-u_i)+\alpha_\pi(\pi)+\varepsilon_{1i}\right)+\left(1-d_i^*(\pi)\right)\left(m_\pi^*(x_i-u_i)+\varepsilon_{0i}\right)$$

in both $\widehat{m}_+(\pi)$ and $\widehat{m}_-(\pi)$. For $\mathfrak{p}=0,\cdots,p$, we can show

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)\left(1-d_i^*(\pi)\right)k_h\left(x_i-\pi\right)m_\pi^*(x_i-u_i)$$

$$\xrightarrow{p}\gamma_{\mathfrak{p}}^+\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*}(x^*)f_u(\pi-x^*)dx^*=\gamma_{\mathfrak{p}}^+f_x(\pi)\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*,$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)\left(1-d_i^*(\pi)\right)k_h\left(x_i-\pi\right)m_\pi^*(x_i-u_i)$$

$$\xrightarrow{p}\gamma_{\mathfrak{p}}^+\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*}(x^*)f_u(\pi-x^*)dx^*=\gamma_{\mathfrak{p}}^+f_x(\pi)\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*,$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)\left[d_i^*(\pi)\varepsilon_{1i}+\left(1-d_i^*(\pi)\right)\varepsilon_{0i}\right]\xrightarrow{p}0,$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}\left(1-d_i(\pi)\right)\left[d_i^*(\pi)\varepsilon_{1i}+\left(1-d_i^*(\pi)\right)\varepsilon_{0i}\right]\xrightarrow{p}0.$$

The limits of the denominators of $\widehat{m}_+(\pi)$ and $\widehat{m}_-(\pi)$ are shown in the proof of Theorem 1. So combining all the limits above, we have

$$\begin{aligned}\widehat{\alpha}(\pi)&\xrightarrow{p}e_1'\Gamma_+^{-1}\mu_{0,p}^+\left(\int_{\pi}^{\infty}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*+\alpha_\pi\left(\pi\right)\left(1-F_{x^*|x}(\pi|\pi)\right)+\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*\right)\\&-e_1'\Gamma_-^{-1}\mu_{0,p}^-\left(\int_{\pi}^{\infty}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*+\alpha_\pi\left(\pi\right)\left(1-F_{x^*|x}(\pi|\pi)\right)+\int_{-\infty}^{\pi}m_\pi^*(x^*)f_{x^*|x}(x^*|\pi)dx^*\right)\\&=\quad 0.\end{aligned}$$

■

Proof of Theorem 3. The proof is quite close to that of Theorem 1. The bias

$$B_{1n}=e_1'S_{n+}^{-1}(\pi)\overline{r}_+\left(\pi\right)-e_1'S_{n-}^{-1}(\pi)\overline{r}_-\left(\pi\right)-\alpha_\pi\left(\pi\right).\tag{17}$$

As in Theorem 1, we show $B_{1n}=o_p(1)$ as follows.

For $\mathfrak{p} = 0, \cdots, p$,

$$
\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)k_h\left(x_i-\pi\right)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right)\right] \\
= & \int\int_0^M k(v)v^{\mathfrak{p}}\left(m_\pi(\pi+vh-n^{-\delta}\epsilon)+\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)\right)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
\to & \gamma_{\mathfrak{p}}^{+}f_{x^*}(\pi)\left[m_\pi(\pi)+\alpha_\pi\left(\pi\right)\right],
\end{aligned}
$$

where the first equality is obtained by changing variables, and the convergence is from the DCT and Assumption F′ and M.

$$
\begin{aligned}
& Var\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)k_h\left(x_i-\pi\right)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right)\right) \\
\leq & \frac{C}{nh}\int\int_0^M k^2(v)v^{2\mathfrak{p}}\left[m_\pi^2(\pi+vh-n^{-\delta}\epsilon)+\alpha_\pi^2(\pi+vh-n^{-\delta}\epsilon)\right]f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
= & O\left(\frac{1}{nh}\right)=o(1).
\end{aligned}
$$

by the DCT. So $\overline{r}_+(\pi) = f_{x^*}(\pi)m_\pi(\pi)\mu_{0,p}^{+} + o_p(1)$. Similarly, $\overline{r}_-(\pi) = f_{x^*}(\pi)m_\pi(\pi)\mu_{0,p}^{-} + o_p(1)$. Turn to the denominator term S_{n+}. For $l = 0, \cdots, 2p$,

$$
E\left[\frac{1}{n}\sum_{i=1}^{n}k_h\left(x_i-\pi\right)d_i(\pi)\left(\frac{x_i-\pi}{h}\right)^{l}\right]=\int\int_0^M k(v)v^l f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon\to f_{x^*}(\pi)\gamma_l^{+}
$$

and

$$
Var\left(\frac{1}{n}\sum_{i=1}^{n}k_h\left(x_i-\pi\right)d_i(\pi)\left(\frac{x_i-\pi}{h}\right)^{l}\right)\leq\frac{1}{nh}\int\int_0^M k^2(v)v^{2l}f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon\to 0
$$

by the DCT. So $S_{n+} = f_{x^*}(\pi)\Gamma_+ + o_p(1)$. By the continuity of matrix inversion, $S_{n+}^{-1} \xrightarrow{p} f_{x^*}^{-1}(\pi)\Gamma_+^{-1}$. Similarly, $S_{n-}^{-1} \xrightarrow{p} f_{x^*}^{-1}(\pi)\Gamma_-^{-1}$. So

$$
B_{1n}\xrightarrow{p}e_1'\Gamma_+^{-1}\mu_{0,p}^{+}\left[m_\pi(\pi)+\alpha_\pi\left(\pi\right)\right]-e_1'\Gamma_-^{-1}\mu_{0,p}^{-}m_\pi(\pi)-\alpha_\pi\left(\pi\right)=0.
$$

We now calculate the asymptotic variance. For $l = 0, \cdots, 2p$,

$$
\begin{aligned}
& \frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^{l}d_i(\pi)\varepsilon_{1i}^2\right] \\
= & \int_0^M\int k^2\left(v\right)v^l\varepsilon^2\left[\int f_{x^*,\varepsilon}^{1}\left(\pi+vh-n^{-\delta}\epsilon,\varepsilon\right)f_\epsilon(\epsilon)d\epsilon\right]d\varepsilon dv \\
= & \int_0^M k^2\left(v\right)v^l\left[\int\sigma_1^2(\pi+vh-n^{-\delta}\epsilon)f_{x^*}\left(\pi+vh-n^{-\delta}\epsilon\right)f_\epsilon(\epsilon)d\epsilon\right]dv \\
\to & \int_0^M k^2\left(v\right)v^l dv\sigma_1^2(\pi)f_{x^*}\left(\pi\right)=f_{x^*}\left(\pi\right)\sigma_1^2(\pi)\omega_l^{+}.
\end{aligned}
$$

The Liapunov's CLT and the Cramer-Wold device can still apply to find that

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}Z_i(\pi)\,d_i(\pi)k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i \xrightarrow{d} N\left(0, f_{x^*}(\pi)\,\sigma_1^2(\pi)\Omega_+\right).$$

Simiarly,

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}Z_i(\pi)\,(1-d_i(\pi))\,k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i \xrightarrow{d} N\left(0, f_{x^*}(\pi)\,E\left[\varepsilon_0^2|x^*=\pi\right]\Omega_-\right).$$

So the asymptotic distribution of $\sqrt{nh}\,(\widehat{\alpha}(\pi)-\alpha_\pi(\pi)-B_{1n})$ follows by Slutsky's theorem. ■

Proof of Theorem 4. In this case, the bias

$$B_{2n}=e_1'\underline{S}_{n+}^{-1}(\pi)\overline{\underline{r}}_+(\pi)-e_1'\underline{S}_{n-}^{-1}(\pi)\overline{\underline{r}}_-(\pi)-\alpha_\pi(\pi), \tag{18}$$

where

$$\begin{aligned}
\overline{\underline{r}}_+(\pi) &= \frac{1}{n}\sum_{i=1}^{n}Z_i(\pi)\,d_i(\pi)d_i^*(\pi)k_h(x_i-\pi)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right),\\
\overline{\underline{r}}_-(\pi) &= \frac{1}{n}\sum_{i=1}^{n}Z_i(\pi)\,(1-d_i(\pi))\,(1-d_i^*(\pi))\,k_h(x_i-\pi)\,m_\pi(x_i-u_i),
\end{aligned}$$

and $\underline{S}_{n+}(\pi)$ and $\underline{S}_{n-}(\pi)$ are defined in Theorem 2.

For $\mathfrak{p}=0,\cdots,p$,

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}d_i(\pi)d_i^*(\pi)k_h(x_i-\pi)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right)\right]\\
=& \int_0^\infty\int_{\frac{n^{-\delta}}{h}\epsilon}^{M}k(v)v^{\mathfrak{p}}\left(m_\pi(\pi+vh-n^{-\delta}\epsilon)+\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)\right)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon\\
&+\int_{-\infty}^{0}\int_0^{M}k(v)v^{\mathfrak{p}}\left(m_\pi(\pi+vh-n^{-\delta}\epsilon)+\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)-\alpha_\pi(\pi)\right)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon\\
\to\ & f_{x^*}(\pi)\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]\begin{cases}\gamma_{\mathfrak{p}}^+, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\gamma}_{\mathfrak{p}}^+(C)+\gamma_{\mathfrak{p}}^+F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \gamma_{\mathfrak{p}}^+F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty,\end{cases}
\end{aligned}$$

by the DCT and Assumptions F′ and M.

$$\begin{aligned}
& Var\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}d_i(\pi)d_i^*(\pi)k_h(x_i-\pi)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right)\right)\\
\le\ & \frac{C}{nh}\int\int_{-M}^{M}k^2(v)v^{2\mathfrak{p}}\left[m_\pi^2(\pi+vh-n^{-\delta}\epsilon)+\alpha_\pi^2(\pi+vh-n^{-\delta}\epsilon)\right]f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon\\
=\ & O\left(\frac{1}{nh}\right)=o(1).
\end{aligned}$$

by Assumptions B, F′ and M. So

$$\underline{\overline{r}}_{+}(\pi) \xrightarrow{p} f_{x^*}(\pi)\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]\begin{cases} \mu_{0,p}^{+}, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\mu}_{0,p}^{+}(C)+\mu_{0,p}^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \mu_{0,p}^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

Similarly,

$$\underline{\overline{r}}_{-}(\pi) \xrightarrow{p} f_{x^*}(\pi)m_\pi(\pi)\begin{cases} \mu_{0,p}^{-}, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\mu}_{0,p}^{-}(C)+\mu_{0,p}^{-}\left(1-F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \mu_{0,p}^{+}\left(1-F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

Turn to the denominator $\underline{S}_{n+}$. For $l=0,\cdots,2p$,

$$E\left[\frac{1}{n}\sum_{i=1}^{n}k_h(x_i-\pi)d_i(\pi)d_i^*(\pi)\left(\frac{x_i-\pi}{h}\right)^l\right]\to f_{x^*}(\pi)\begin{cases} \gamma_l^{+}, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\gamma}_l^{+}(C)+\gamma_l^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \gamma_l^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty, \end{cases}$$

and

$$Var\left(\frac{1}{n}\sum_{i=1}^{n}k_h(x_i-\pi)d_i(\pi)d_i^*(\pi)\left(\frac{x_i-\pi}{h}\right)^l\right)\to 0$$

by the DCT. So

$$\underline{S}_{n+}\xrightarrow{p} f_{x^*}(\pi)\begin{cases} \Gamma_{+}, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\Gamma}_{+}(C)+\Gamma_{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \Gamma_{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

By the continuity of matrix inversion,

$$\underline{S}_{n+}^{-1}\xrightarrow{p} f_{x^*}^{-1}(\pi)\begin{cases} \Gamma_{+}^{-1}+o_p(1), & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \left[\overline{\Gamma}_{+}(C)+\Gamma_{+}F_\epsilon(0)\right]^{-1}+o_p(1), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \Gamma_{+}^{-1}F_\epsilon(0)^{-1}+o_p(1), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

Similarly,

$$\underline{S}_{n-}^{-1}\xrightarrow{p} f_{x^*}^{-1}(\pi)\begin{cases} \Gamma_{-}^{-1}, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \left[\overline{\Gamma}_{-}(C)+\Gamma_{-}\left(1-F_\epsilon(0)\right)\right]^{-1}, & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \Gamma_{-}^{-1}\left(1-F_\epsilon(0)\right)^{-1}, & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}.$$

In summary,

$$B_{2n}\xrightarrow{p}\begin{cases} e_1'\Gamma_{+}^{-1}\mu_{0,p}^{+}\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]-e_1'\Gamma_{-}^{-1}\mu_{0,p}^{-}m_\pi(\pi)-\alpha_\pi(\pi)=0, \text{ if } \frac{n^{-\delta}}{h}\to 0,\\ e_1'\left[\overline{\Gamma}_{+}(C)+\Gamma_{+}F_\epsilon(0)\right]^{-1}\left[\overline{\mu}_{0,p}^{+}(C)+\mu_{0,p}^{+}F_\epsilon(0)\right]\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]\\ -e_1'\left[\overline{\Gamma}_{-}(C)+\Gamma_{-}\left(1-F_\epsilon(0)\right)\right]^{-1}\left[\overline{\mu}_{0,p}^{-}(C)+\mu_{0,p}^{-}\left(1-F_\epsilon(0)\right)\right]m_\pi(\pi)\\ -\alpha_\pi(\pi)=0, \quad \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ e_1'\Gamma_{+}^{-1}\mu_{0,p}^{+}\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]-e_1'\Gamma_{-}^{-1}\mu_{0,p}^{-}m_\pi(\pi)-\alpha_\pi(\pi)=0, \text{ if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

The Liapunov's CLT and the Cramer-Wold device can still apply to find the asymptotic distribution of

$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi) d_i^*(\pi) k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i$. First note that for $U_{ni} = k\left(\frac{x_i-\pi}{h}\right) d_i(\pi)d_i^*(\pi)\lambda' Z_i(\pi)\,\varepsilon_i/\sqrt{nh}$,

$$\begin{aligned} E[U_{ni}] &= E[E[U_{ni}|x_i, x_i^*]] \\ &= E\left[k\left(\frac{x_i-\pi}{h}\right) d_i(\pi)d_i^*(\pi)\lambda' Z_i(\pi)/\sqrt{nh} E\left[\varepsilon_i|u_i, x_i^*\right]\right] \\ &= E\left[k\left(\frac{x_i-\pi}{h}\right) d_i(\pi)d_i^*(\pi)\lambda' Z_i(\pi)/\sqrt{nh} E\left[\varepsilon_i|x_i^*\right]\right] = 0. \end{aligned}$$

We then calculate the asymptotic variance as follows. For $l = 0, \cdots, 2p$,

$$\begin{aligned} & \frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi)d_i^*(\pi)\varepsilon_i^2\right] \\ = & \int_0^\infty \int_{\frac{n^{-\delta}}{h}\epsilon}^M \int k^2(v)\, v^l \varepsilon^2 f^1_{x^*,\varepsilon}\left(\pi + vh - n^{-\delta}\epsilon, \varepsilon\right) f_\epsilon(\epsilon) d\varepsilon dv d\epsilon \\ & + \int_{-\infty}^0 \int_0^M \int k^2(v)\, v^l \varepsilon^2 f^1_{x^*,\varepsilon}\left(\pi + vh - n^{-\delta}\epsilon, \varepsilon\right) f_\epsilon(\epsilon) d\varepsilon dv d\epsilon \\ \to & \; f_{x^*}(\pi)\,\sigma_1^2(\pi) \begin{cases} \omega_l^+, & \text{if } \frac{n^{-\delta}}{h} \to 0, \\ \overline{\omega}_l^+(C) + \omega_l^+ F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h} \to C \in (0,\infty), \\ \omega_l^+ F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h} \to \infty. \end{cases} \end{aligned}$$

So

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi)d_i^*(\pi) k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i \xrightarrow{d} N\left(0, f_{x^*}(\pi)\,\sigma_1^2(\pi)\widetilde{\Omega}_+\right),$$

where

$$\widetilde{\Omega}_+ = \begin{cases} \Omega_+, & \text{if } \frac{n^{-\delta}}{h} \to 0, \\ \overline{\Omega}_+(C) + \Omega_+ F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h} \to C \in (0,\infty), \\ \Omega_+ F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h} \to \infty. \end{cases}$$

Similarly,

$$\frac{1}{\sqrt{nh}}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi)\left(1-d_i(\pi)\right)\left(1-d_i^*(\pi)\right) k\left(\frac{x_i-\pi}{h}\right)\varepsilon_i \xrightarrow{d} N\left(0, f_{x^*}(\pi)\,\sigma_0^2(\pi)\widetilde{\Omega}_-\right),$$

where

$$\widetilde{\Omega}_+ = \begin{cases} \Omega_-, & \text{if } \frac{n^{-\delta}}{h} \to 0, \\ \overline{\Omega}_-(C) + \Omega_-\left(1 - F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h} \to C \in (0,\infty), \\ \Omega_-\left(1 - F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h} \to \infty. \end{cases}$$

So the asymptotic distribution of $\sqrt{nh}\left(\widehat{\alpha}(\pi) - \alpha_\pi(\pi) - B_{2n}\right)$ follows by Slutsky's theorem. ■

Proof of Theorem 5. The estimator is the same as (14), but now y_i in $\widehat{m}_+(\pi)$ contains also controlled individuals, and y_i in $\widehat{m}_-(\pi)$ contains also treated individuals. In this case,

$$B_{3n} = e_1' S_{n+}^{-1}(\pi)\overline{\overline{r}}_+(\pi) - e_1' S_{n-}^{-1}(\pi)\overline{\overline{r}}_-(\pi) - \alpha_\pi(\pi), \qquad (19)$$

where

$$\begin{aligned}
\overline{\overline{r}}_{+}(\pi) &= \frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi) d_i^*(\pi) k_h(x_i-\pi)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi)\left(1-d_i^*(\pi)\right) k_h(x_i-\pi)\, m_\pi(x_i-u_i) \equiv \overline{\overline{r}}_{+}^{t}(\pi) + \overline{\overline{r}}_{+}^{c}(\pi), \\
\overline{\overline{r}}_{-}(\pi) &= \frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\left(1-d_i(\pi)\right) d_i^*(\pi) k_h(x_i-\pi)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right) \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\left(1-d_i(\pi)\right)\left(1-d_i^*(\pi)\right) k_h(x_i-\pi)\, m_\pi(x_i-u_i) \equiv \overline{\overline{r}}_{-}^{t}(\pi) + \overline{\overline{r}}_{-}^{c}(\pi).
\end{aligned}$$

As in the proof of Theorem 4,

$$\overline{\overline{r}}_{+}^{t}(\pi) \xrightarrow{p} f_{x^*}(\pi)\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]\begin{cases} \mu_{0,p}^{+}, & \text{if } \frac{n^{-\delta}}{h}\to 0, \\ \overline{\mu}_{0,p}^{+}(C)+\mu_{0,p}^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty), \\ \mu_{0,p}^{+}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

$$\overline{\overline{r}}_{+}^{c}(\pi) \xrightarrow{p} f_{x^*}(\pi) m_\pi(\pi)\begin{cases} 0, & \text{if } \frac{n^{-\delta}}{h}\to 0, \\ \mu_{0,p}^{+}\left(1-F_\epsilon(0)\right)-\overline{\mu}_{0,p}^{+}(C), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty), \\ \mu_{0,p}^{+}\left(1-F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h}\to\infty, \end{cases}$$

and

$$\overline{\overline{r}}_{-}^{t}(\pi) \xrightarrow{p} f_{x^*}(\pi)\left[m_\pi(\pi)+\alpha_\pi(\pi)\right]\begin{cases} 0, & \text{if } \frac{n^{-\delta}}{h}\to 0, \\ \mu_{0,p}^{-}F_\epsilon(0)-\overline{\mu}_{0,p}^{-}(C), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty), \\ \mu_{0,p}^{-}F_\epsilon(0), & \text{if } \frac{n^{-\delta}}{h}\to\infty, \end{cases}$$

$$\overline{\overline{r}}_{-}^{c}(\pi) \xrightarrow{p} f_{x^*}(\pi)\, m_\pi(\pi)\begin{cases} \mu_{0,p}^{-}, & \text{if } \frac{n^{-\delta}}{h}\to 0, \\ \overline{\mu}_{0,p}^{-}(C)+\mu_{0,p}^{-}\left(1-F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty), \\ \mu_{0,p}^{-}\left(1-F_\epsilon(0)\right), & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

The limit of the denominator is the same as in Theorem 2:

$$S_{n+}^{-1} \xrightarrow{p} f_{x^*}^{-1}(\pi)\Gamma_{+}^{-1},\ S_{n-}^{-1} \xrightarrow{p} f_{x^*}^{-1}(\pi)\Gamma_{-}^{-1}.$$

In whatever cases,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi) k_h(x_i-\pi)\,\varepsilon_i &= o_p(1), \\
\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\left(1-d_i(\pi)\right) k_h(x_i-\pi)\,\varepsilon_i &= o_p(1).
\end{aligned}$$

In summary,

$$\widehat{\alpha}(\pi) \xrightarrow{p} \begin{cases} [m_\pi(\pi)+\alpha_\pi(\pi)]\, e_1'\Gamma_+^{-1}\mu_{0,p}^+ - m_\pi(\pi)e_1'\Gamma_-^{-1}\mu_{0,p}^- = \alpha_\pi(\pi), & \text{if } \frac{n^{-\delta}}{h}\to 0, \\ [m_\pi(\pi)+\alpha_\pi(\pi)]\left[e_1'\Gamma_+^{-1}\left(\overline{\mu}_{0,p}^+(C)+\mu_{0,p}^+F_\epsilon(0)\right) - e_1'\Gamma_-^{-1}\left(\mu_{0,p}^-F_\epsilon(0)-\overline{\mu}_{0,p}^-(C)\right)\right] & \\ +m_\pi(\pi)\left[e_1'\Gamma_+^{-1}\left(\mu_{0,p}^+(1-F_\epsilon(0))-\overline{\mu}_{0,p}^+(C)\right) - e_1'\Gamma_-^{-1}\left(\overline{\mu}_{0,p}^-(C)+\mu_{0,p}^-(1-F_\epsilon(0))\right)\right] & \\ = \alpha_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right], & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty), \\ [m_\pi(\pi)+\alpha_\pi(\pi)]\left[e_1'\Gamma_+^{-1}\mu_{0,p}^+F_\epsilon(0) - e_1'\Gamma_-^{-1}\mu_{0,p}^-F_\epsilon(0)\right] & \\ +m_\pi(\pi)\left[e_1'\Gamma_+^{-1}\mu_{0,p}^+(1-F_\epsilon(0)) - e_1'\Gamma_-^{-1}\mu_{0,p}^-(1-F_\epsilon(0))\right]=0, & \text{if } \frac{n^{-\delta}}{h}\to\infty. \end{cases}$$

When $\frac{n^{-\delta}}{h}\to 0$, the asymptotic variance is the same as in Theorem 3. ■

Proof of Theorem 6. Under Assumption CY and U,

$$\begin{aligned} m_\pi(x_i-u_i) &= m_\pi(\pi)+m_\pi'(\pi)(x_i-\pi-u_i)+\frac{m_\pi''(\pi)}{2}(x_i-\pi-u_i)^2+R_m(x_i-u_i,\pi), \\ \alpha_\pi(x_i-u_i) &= \alpha_\pi(\pi)+\alpha_\pi'(\pi)(x_i-\pi-u_i)+\frac{\alpha_\pi''(\pi)}{2}(x_i-\pi-u_i)^2+R_\alpha(x_i-u_i,\pi), \end{aligned}$$

where $R_m(x_i-u_i,\pi)$ and $R_\alpha(x_i-u_i,\pi)$ are the remaining terms in the second-order Taylor expansion of $m_\pi(x_i-u_i)$ and $\alpha_\pi(x_i-u_i)$ at π, respectively. By the discrete orthogonality relation,

$$\begin{aligned} B_{1n} &\approx e_1'S_{n+}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\left[-(m_\pi'(\pi)+\alpha_\pi'(\pi))\,u_i+\frac{m_\pi''(\pi)+\alpha_\pi''(\pi)}{2}(x_i-\pi-u_i)^2\right] \\ &\quad -e_1'S_{n-}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,(1-d_i(\pi))\,k_h(x_i-\pi)\left[-m_\pi'(\pi)\,u_i+\frac{m_\pi''(\pi)}{2}(x_i-\pi-u_i)^2\right], \end{aligned}$$

where $\approx$ means higher-order terms are omitted. For $\mathfrak{p}=0,1$,

$$\begin{aligned} & E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)k_h(x_i-\pi)\,u_i\right] \\ =\; & n^{-\delta}\int\int_0^M k(v)v^{\mathfrak{p}}\epsilon f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\ \approx\; & f_{x^*}(\pi)n^{-\delta}\gamma_{\mathfrak{p}}^+E[\epsilon]+f_{x^*}'(\pi)\left(hn^{-\delta}\gamma_{\mathfrak{p}+1}^+E[\epsilon]-n^{-2\delta}\gamma_{\mathfrak{p}}^+E\left[\epsilon^2\right]\right), \end{aligned}$$

and

$$\begin{aligned} & E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)k_h(x_i-\pi)(x_i-\pi-u_i)^2\right] \\ =\; & \int\int_0^M k(v)v^{\mathfrak{p}}\left(vh-n^{-\delta}\epsilon\right)^2 f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\ \approx\; & f_{x^*}(\pi)\left(h^2\gamma_{\mathfrak{p}+2}^+-2hn^{-\delta}\gamma_{\mathfrak{p}+1}^+E[\epsilon]+n^{-2\delta}\gamma_{\mathfrak{p}}^+E\left[\epsilon^2\right]\right). \end{aligned}$$

For $l = 0, 1, 2,$

$$\begin{aligned} & E\left[\frac{1}{n}\sum_{i=1}^{n} k_h\left(x_i-\pi\right)d_i(\pi)\left(\frac{x_i-\pi}{h}\right)^l\right] \\ = & \int\int_0^M k(v)v^l f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\ \approx & f_{x^*}(\pi)\gamma_l^+ + f'_{x^*}(\pi)\left(h\gamma_{l+1}^+ - n^{-\delta}\gamma_l^+ E\left[\epsilon\right]\right), \end{aligned}$$

so

$$\begin{aligned} S_{n+}^{-1}(\pi) & \approx \left(f_{x^*}(\pi)\Gamma_+ + hf'_{x^*}(\pi)\Gamma_{(+1)} - n^{-\delta}f'_{x^*}(\pi)E\left[\epsilon\right]\Gamma_+\right)^{-1} \\ & = \frac{1}{f_{x^*}(\pi)}\Gamma_+^{-1} - h\frac{f'_{x^*}(\pi)}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1} + n^{-\delta}\frac{f'_{x^*}(\pi)E\left[\epsilon\right]}{f_{x^*}^2(\pi)}\Gamma_+^{-1}, \end{aligned}$$

where $\Gamma_{(+1)} = \begin{pmatrix} \gamma_1^+ & \gamma_2^+ \\ \gamma_2^+ & \gamma_3^+ \end{pmatrix}$. Similar analysis applies to the second term of B_{1n}. In summary,

$$\begin{aligned} & B_{1n} \approx e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_+^{-1} - h\frac{f'_{x^*}(\pi)}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1} + n^{-\delta}\frac{f'_{x^*}(\pi)E\left[\epsilon\right]}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\right) \\ & \left[\begin{array}{c} -\left(m'_\pi\left(\pi\right)+\alpha'_\pi\left(\pi\right)\right)\left[f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^+E\left[\epsilon\right] + f'_{x^*}(\pi)\left(hn^{-\delta}\mu_{1,2}^+E\left[\epsilon\right] - n^{-2\delta}\mu_{0,1}^+E\left[\epsilon^2\right]\right)\right] \\ +\frac{m''_\pi(\pi)+\alpha''_\pi(\pi)}{2}f_{x^*}(\pi)\left(h^2\mu_{2,3}^+ - 2hn^{-\delta}\mu_{1,2}^+E\left[\epsilon\right] + n^{-2\delta}\mu_{0,1}^+E\left[\epsilon^2\right]\right) \end{array}\right] \\ & -e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_-^{-1} - h\frac{f'_{x^*}(\pi)}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1} + n^{-\delta}\frac{f'_{x^*}(\pi)E\left[\epsilon\right]}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\right) \\ & \left[\begin{array}{c} -m'_\pi\left(\pi\right)\left[f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^-E\left[\epsilon\right] + f'_{x^*}(\pi)\left(hn^{-\delta}\mu_{1,2}^-E\left[\epsilon\right] - n^{-2\delta}\mu_{0,1}^-E\left[\epsilon^2\right]\right)\right] \\ +\frac{m''_\pi(\pi)}{2}f_{x^*}(\pi)\left(h^2\mu_{2,3}^- - 2hn^{-\delta}\mu_{1,2}^-E\left[\epsilon\right] + n^{-2\delta}\mu_{0,1}^-E\left[\epsilon^2\right]\right) \end{array}\right] \\ \approx & -n^{-\delta}\alpha'_\pi\left(\pi\right)E\left[\epsilon\right] + n^{-2\delta}\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\alpha'_\pi\left(\pi\right)E\left[\epsilon^2\right] + h^2\frac{\alpha''_\pi\left(\pi\right)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+ + n^{-2\delta}\frac{\alpha''_\pi\left(\pi\right)}{2}E\left[\epsilon^2\right] \\ & +hn^{-\delta}\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}E\left[\epsilon\right]\left[\left(m'_\pi\left(\pi\right)+\alpha'_\pi\left(\pi\right)\right)e_1'\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1}\mu_{0,1}^+ - m'_\pi\left(\pi\right)e_1'\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1}\mu_{0,1}^-\right] \\ & -n^{-2\delta}\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}E\left[\epsilon\right]^2\alpha'_\pi\left(\pi\right) \\ = & -n^{-\delta}\alpha'_\pi\left(\pi\right)E\left[\epsilon\right] + n^{-2\delta}\left[\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\alpha'_\pi\left(\pi\right)Var(\epsilon) + \frac{\alpha''_\pi\left(\pi\right)}{2}E\left[\epsilon^2\right]\right] + h^2\frac{\alpha''_\pi\left(\pi\right)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+. \end{aligned}$$

where $\Gamma_{(-1)}$ is similarly defined as $\Gamma_{(+1)}$ with γ_j^+ replaced by γ_j^-, and the last equality is from the fact that $e_1'\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1}\mu_{0,1}^+ = e_1'\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1}\mu_{0,1}^- = 0$.

For B_{2n}, we need only consider the first-order Taylor expansion of $m_\pi(x_i - u_i)$ and $\alpha_\pi(x_i - u_i)$:

$$\begin{aligned} B_{2n} \approx & e_1'\underline{S}_{n+}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n} Z_i\left(\pi\right)d_i(\pi)d_i^*(\pi)k_h\left(x_i-\pi\right)\left[-\left(m'_\pi\left(\pi\right)+\alpha'_\pi\left(\pi\right)\right)u_i\right] \\ & -e_1'\underline{S}_{n-}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n} Z_i\left(\pi\right)\left(1-d_i(\pi)\right)\left(1-d_i^*(\pi)\right)\left[-m'_\pi\left(\pi\right)u_i\right]. \end{aligned}$$

For $\mathfrak{p} = 0, 1$,

$$\begin{aligned} & E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)d_i^*(\pi)k_h\left(x_i-\pi\right)u_i\right] \\ = & \int_0^{\infty}\int_{\frac{n^{-\delta}}{h}\epsilon}^{M} k(v)v^{\mathfrak{p}}n^{-\delta}\epsilon f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\ & +\int_{-\infty}^{0}\int_0^{M} k(v)v^{\mathfrak{p}}n^{-\delta}\epsilon f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\ \approx & n^{-\delta}f_{x^*}(\pi)\left[\overline{\overline{\gamma}}_{\mathfrak{p}}^{+}(C)+\gamma_{\mathfrak{p}}^{+}\int_{-\infty}^{0}\epsilon f_\epsilon(\epsilon)d\epsilon\right]. \end{aligned}$$

Similarly, $E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}\left(1-d_i(\pi)\right)\left(1-d_i^*(\pi)\right)k_h\left(x_i-\pi\right)u_i\right] \approx n^{-\delta}f_{x^*}(\pi)\left[\overline{\overline{\gamma}}_{\mathfrak{p}}^{-}(C)+\gamma_{\mathfrak{p}}^{-}\int_0^{\infty}\epsilon f_\epsilon(\epsilon)d\epsilon\right]$.
The limit of the denominator is derived in Theorem 4. So

$$\begin{aligned} B_{2n} \approx & -n^{-\delta}e_1'\frac{1}{f_{x^*}(\pi)}\left[\overline{\Gamma}_+(C)+\Gamma_+F_\epsilon(0)\right]^{-1}\left[m_\pi'(\pi)+\alpha_\pi'(\pi)\right]f_{x^*}(\pi)\left(\overline{\overline{\mu}}_{0,1}^{+}(C)+\mu_{0,1}^{+}E[\epsilon|\epsilon<0]F_\epsilon(0)\right) \\ & -\left(-n^{-\delta}\right)e_1'\frac{1}{f_{x^*}(\pi)}\left[\overline{\Gamma}_-(C)+\Gamma_-\left(1-F_\epsilon(0)\right)\right]^{-1}m_\pi'(\pi)f_{x^*}(\pi)\left(\overline{\overline{\mu}}_{0,1}^{-}(C)+\mu_{0,1}^{-}E[\epsilon|\epsilon>0]\left(1-F_\epsilon(0)\right)\right) \\ = & -n^{-\delta}\left\{\left(m_\pi'(\pi)+\alpha_\pi'(\pi)\right)e_1'\left[\overline{\Gamma}_+(C)+\Gamma_+F_\epsilon(0)\right]^{-1}\left[\overline{\overline{\mu}}_{0,1}^{+}(C)+\mu_{0,1}^{+}E[\epsilon|\epsilon<0]F_\epsilon(0)\right]\right. \\ & \left.-m_\pi'(\pi)e_1'\left[\overline{\Gamma}_-(C)+\Gamma_-\left(1-F_\epsilon(0)\right)\right]^{-1}\left[\overline{\overline{\mu}}_{0,1}^{-}(C)+\mu_{0,1}^{-}E[\epsilon|\epsilon>0]\left(1-F_\epsilon(0)\right)\right]\right\}. \end{aligned}$$

■

Proof of Theorem 7. The estimator $\widehat{\alpha}_f(\pi)$ is defined in (6). In the fuzzy design,

$$y_i = D_i\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)+\varepsilon_{1i}\right)+\left(1-D_i\right)\left(m_\pi(x_i-u_i)+\varepsilon_{0i}\right).$$

When the forcing variable is x,

$$\begin{aligned} y_i = & \left(p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)+\eta_i\right)\left(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)+\varepsilon_{1i}\right) \\ & +\left(1-p_\pi(x_i)-\beta_\pi(x_i)d_i(\pi)-\eta_i\right)\left(m_\pi(x_i-u_i)+\varepsilon_{0i}\right) \\ = & \left[p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)\right]\left[m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right] \\ & +\left[1-p_\pi(x_i)-\beta_\pi(x_i)d_i(\pi)\right]m_\pi(x_i-u_i) \\ & +\left[p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)\right]\varepsilon_{1i}+\left[1-p_\pi(x_i)-\beta_\pi(x_i)d_i(\pi)\right]\varepsilon_{0i} \\ & +\alpha_\pi(x_i-u_i)\eta_i+\eta_i\left(\varepsilon_{1i}-\varepsilon_{0i}\right), \\ \equiv & \overline{y}_i+R_i \end{aligned}$$

where

$$\begin{aligned} \overline{y}_i = & \left[p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)\right]\left[m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)\right] \\ & +\left[1-p_\pi(x_i)-\beta_\pi(x_i)d_i(\pi)\right]m_\pi(x_i-u_i) \\ = & \beta_\pi(x_i)\alpha_\pi(x_i-u_i)d_i(\pi)+m_\pi(x_i-u_i)+p_\pi(x_i)\alpha_\pi(x_i-u_i), \\ R_i = & y_i-\overline{y}_i; \end{aligned}$$

when the forcing variable is x^*,

$$\begin{aligned} y_i &= (p_\pi(x_i-u_i)+\beta_\pi(x_i-u_i)d_i^*(\pi)+\eta_i)\,(m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)+\varepsilon_{1i}) \\ &\quad +(1-p_\pi(x_i-u_i)-\beta_\pi(x_i-u_i)d_i^*(\pi)-\eta_i)\,(m_\pi(x_i-u_i)+\varepsilon_{0i}) \\ &= [p_\pi(x_i-u_i)+\beta_\pi(x_i-u_i)d_i^*(\pi)]\,[m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)] \\ &\quad +[1-p_\pi(x_i-u_i)-\beta_\pi(x_i-u_i)d_i^*(\pi)]\,m_\pi(x_i-u_i) \\ &\quad +[p_\pi(x_i^*)+\beta_\pi(x_i^*)d_i^*(\pi)]\,\varepsilon_{1i}++[1-p_\pi(x_i^*)-\beta_\pi(x_i^*)d_i^*(\pi)]\,\varepsilon_{0i} \\ &\quad +\alpha_\pi(x_i^*)\eta_i+\eta_i\,(\varepsilon_{1i}-\varepsilon_{0i}) \\ &\equiv \overline{y}_i^*+R_i^*, \end{aligned}$$

where

$$\begin{aligned} \overline{y}_i^* &= [p_\pi(x_i-u_i)+\beta_\pi(x_i-u_i)d_i^*(\pi)]\,[m_\pi(x_i-u_i)+\alpha_\pi(x_i-u_i)] \\ &\quad +[1-p_\pi(x_i-u_i)-\beta_\pi(x_i-u_i)d_i^*(\pi)]\,m_\pi(x_i-u_i) \\ &= \beta_\pi(x_i-u_i)\alpha_\pi(x_i-u_i)d_i^*(\pi)+m_\pi(x_i-u_i)+p_\pi(x_i-u_i)\alpha_\pi(x_i-u_i), \\ R_i^* &= y_i^*-\overline{y}_i^*. \end{aligned}$$

We analyze the four cases in sequence. First define some notations,

$$\begin{aligned} B_m^+ &= e_1'S_{n+}^{-1}(\pi)\left(\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\,\overline{y}_i\right), \\ V_m^+ &= e_1'S_{n+}^{-1}(\pi)\left(\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\,R_i\right), \\ B_p^+ &= e_1'S_{n+}^{-1}(\pi)\left(\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\,\overline{D}_i\right), \\ B_p^+ &= e_1'S_{n+}^{-1}(\pi)\left(\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\,\eta_i\right), \end{aligned}$$

where $\overline{D}_i = p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)$. B_m^-, V_m^-, B_p^- and V_p^- are similarly defined as B_m^+, V_m^+, B_p^+, V_p^+ but replacing $d_i(\pi)$ by $1-d_i(\pi)$. B_m^{*+}, V_m^{*+}, B_p^{*+}, V_p^{*+}, B_m^{*-}, V_m^{*-}, B_p^{*-} and V_p^{*-} are similarly defined as B_m^+, V_m^+, B_p^+, V_p^+, B_m^-, V_m^-, B_p^- and V_p^- but replacing $d_i(\pi)$, $\overline{y}_i$, R_i and $\overline{D}_i$ by $d_i^*(\pi)$, $\overline{y}_i^*$, R_i^* and $\overline{D}_i^* \equiv p_\pi(x_i^*)+\beta_\pi(x_i^*)d_i^*(\pi)$, respectively. Now, in Case 4,

$$\widehat{\alpha}_f(\pi) = \frac{(B_m^+-B_m^-)+(V_m^+-V_m^-)}{\left(B_p^+-B_p^-\right)+\left(V_p^+-V_p^-\right)},$$

and in Case 5,

$$\widehat{\alpha}_f(\pi) = \frac{(B_m^{*+}-B_m^{*-})+(V_m^{*+}-V_m^{*-})}{\left(B_p^{*+}-B_p^{*-}\right)+\left(V_p^{*+}-V_p^{*-}\right)}.$$

In Case (i) and (ii), the limits of the denominators $S_{n+}(\pi)$ and $S_{n-}(\pi)$ are the same as in Theorem 1. In Case (iii) and (iv), the limits of the denominators are the same as in Theorem 3. So we concentrate on the numerators in each case.

Case (i). Define

$$B_n^f = \frac{B_m^+ - B_m^-}{B_p^+ - B_p^-} - \overline{\alpha}_\pi(\pi), \tag{20}$$

then

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \overline{\alpha}_\pi(\pi) - B_n^f\right) = \frac{\left(B_p^+ - B_p^-\right)\sqrt{nh}\left(V_m^+ - V_m^-\right) - \left(B_m^+ - B_m^-\right)\sqrt{nh}\left(V_p^+ - V_p^-\right)}{\left(B_p^+ - B_p^-\right)\left(B_p^+ - B_p^- + V_p^+ - V_p^-\right)}.$$

For $\mathfrak{p} = 0, \cdots, p$,

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} d_i(\pi) k_h(x_i - \pi)\overline{y}_i\right] \\
= & \int\int_0^M k(v)v^{\mathfrak{p}}\beta_\pi(\pi + vh)\alpha_\pi(\pi + vh - u) f_{x^*}(\pi + vh - u) f_u(u) dv du \\
& + \int\int_0^M k(v)v^{\mathfrak{p}}\left[m_\pi(\pi + vh - u) + p_\pi(\pi + vh)\alpha_\pi(\pi + vh - u)\right] f_{x^*}(\pi + vh - u) f_u(u) dv du \\
\to & \gamma_{\mathfrak{p}}^+ f_x(\pi)\left\{\left[p_\pi(\pi) + \beta_\pi(\pi)\right] E\left[\alpha_\pi(\pi - u)|x = \pi\right] + E\left[m_\pi(\pi - u)|x = \pi\right]\right\},
\end{aligned}$$

and $Var\left(\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} d_i(\pi) k_h(x_i - \pi)\overline{y}_i\right) \to 0$ by the DCT. Similarly,

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} \left(1 - d_i(\pi)\right) k_h(x_i - \pi)\overline{y}_i \\
& \xrightarrow{p} \gamma_{\mathfrak{p}}^- f_x(\pi)\left\{p_\pi(\pi) E\left[\alpha_\pi(\pi - u)|x = \pi\right] + E\left[m_\pi(\pi - u)|x = \pi\right]\right\}.
\end{aligned}$$

So $B_m^+ - B_m^- \xrightarrow{p} e_1'\Gamma_+^{-1}\mu_{0,p}^+\left[\overline{m}_\pi(\pi) + \left(p_\pi(\pi) + \beta_\pi(\pi)\right)\overline{\alpha}_\pi(\pi)\right] - e_1'\Gamma_-^{-1}\mu_{0,p}^-\left[\overline{m}_\pi(\pi) + p_\pi(\pi)\overline{\alpha}_\pi(\pi)\right] = \beta_\pi(\pi)\overline{\alpha}_\pi(\pi)$. Similarly, $B_p^+ - B_p^- \xrightarrow{p} \beta_\pi(\pi)$. So $B_n^f = o_p(1)$.

Next, we derive the joint asymptotic distribution of $\left(\sqrt{nh}\left(V_m^+ - V_m^-\right), \sqrt{nh}\left(V_p^+ - V_p^-\right)\right)$.

$$\begin{aligned}
\begin{pmatrix} \sqrt{nh}\left(V_m^+ - V_m^-\right) \\ \sqrt{nh}\left(V_p^+ - V_p^-\right) \end{pmatrix} &= \begin{pmatrix} e_1'S_{n+}^{-1}(\pi) & 0 \\ 0 & e_1'S_{n+}^{-1}(\pi) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi) d_i(\pi) k\left(\frac{x_i - \pi}{h}\right) R_i \\ \frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi) d_i(\pi) k\left(\frac{x_i - \pi}{h}\right)\eta_i \end{pmatrix} \\
& - \begin{pmatrix} e_1'S_{n-}^{-1}(\pi) & 0 \\ 0 & e_1'S_{n-}^{-1}(\pi) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi)\left(1 - d_i(\pi)\right) k\left(\frac{x_i - \pi}{h}\right) R_i \\ \frac{1}{\sqrt{nh}}\sum_{i=1}^n Z_i(\pi)\left(1 - d_i(\pi)\right) k\left(\frac{x_i - \pi}{h}\right)\eta_i \end{pmatrix} \\
&\equiv \begin{pmatrix} e_1'S_{n+}^{-1}(\pi) & 0 \\ 0 & e_1'S_{n+}^{-1}(\pi) \end{pmatrix} A_+ - \begin{pmatrix} e_1'S_{n-}^{-1}(\pi) & 0 \\ 0 & e_1'S_{n-}^{-1}(\pi) \end{pmatrix} A_-.
\end{aligned}$$

From Assumption ED, $E\left[\eta\left(\varepsilon_1 - \varepsilon_0\right)|x\right] = 0$ and $E[\eta|x] = 0$, so the means of A_+ and A_- are zero. We calculate their variances next. For $l = 0, \cdots, 2p$,

$$\begin{aligned}
& \frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi)R_i^2\right] \\
= & \frac{1}{h}E\left[k^2\left(\frac{x-\pi}{h}\right)\left(\frac{x-\pi}{h}\right)^l d_\pi E\left[R^2|x,u\right]\right] \\
= & \int_0^M k^2(v)\, v^l \int E\left[R^2|x=\pi+vh,u\right] f_{x^*}(\pi+vh-u)\, f_u(u)dudv \\
\to & f_x(\pi)\, E\left[R^2|x=\pi+\right]\omega_l^+,
\end{aligned}$$

and similarly,

$$\begin{aligned}
\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi)\eta_i^2\right] & \to f_x(\pi)\, E\left[\eta^2|x=\pi+\right]\omega_l^+, \\
\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi)R_i\eta_i\right] & \to f_x(\pi)\, E\left[R\eta|x=\pi+\right]\omega_l^+, \\
\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l (1-d_i(\pi))\, R_i^2\right] & \to f_x(\pi)\, E\left[R^2|x=\pi-\right]\omega_l^-, \\
\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l (1-d_i(\pi))\, \eta_i^2\right] & \to f_x(\pi)\, E\left[\eta^2|x=\pi-\right]\omega_l^-, \\
\frac{1}{h}E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l (1-d_i(\pi))\, R_i\eta_i\right] & \to f_x(\pi)\, E\left[R\eta|x=\pi-\right]\omega_l^-.
\end{aligned}$$

So A_+ and A_- are independent, and

$$\begin{aligned}
A_+ & \xrightarrow{d} N\left(0, f_x(\pi)\begin{pmatrix} E\left[R^2|x=\pi+\right] & E\left[R\eta|x=\pi+\right] \\ E\left[R\eta|x=\pi+\right] & E\left[\eta^2|x=\pi+\right]\end{pmatrix}\otimes\Omega_+\right), \\
A_- & \xrightarrow{d} N\left(0, f_x(\pi)\begin{pmatrix} E\left[R^2|x=\pi-\right] & E\left[R\eta|x=\pi-\right] \\ E\left[R\eta|x=\pi-\right] & E\left[\eta^2|x=\pi-\right]\end{pmatrix}\otimes\Omega_-\right),
\end{aligned}$$

where $\otimes$ is the Kronecker product. As a result,

$$\begin{pmatrix} \sqrt{nh}\left(V_m^+ - V_m^-\right) \\ \sqrt{nh}\left(V_p^+ - V_p^-\right)\end{pmatrix} \xrightarrow{d} N\left(0, \Sigma_{mp}\right),$$

where

$$\begin{aligned}
\Sigma_{mp} = & \frac{1}{f_x(\pi)}\begin{pmatrix} e_1'\Gamma_+^{-1} & 0 \\ 0 & e_1'\Gamma_+^{-1}\end{pmatrix}\left[\begin{pmatrix} E\left[R^2|x=\pi+\right] & E\left[R\eta|x=\pi+\right] \\ E\left[R\eta|x=\pi+\right] & E\left[\eta^2|x=\pi+\right]\end{pmatrix}\otimes\Omega_+\right]\begin{pmatrix}\Gamma_+^{-1}e_1 & 0 \\ 0 & \Gamma_+^{-1}e_1\end{pmatrix} \\
& + \frac{1}{f_x(\pi)}\begin{pmatrix} e_1'\Gamma_-^{-1} & 0 \\ 0 & e_1'\Gamma_-^{-1}\end{pmatrix}\left[\begin{pmatrix} E\left[R^2|x=\pi-\right] & E\left[R\eta|x=\pi-\right] \\ E\left[R\eta|x=\pi-\right] & E\left[\eta^2|x=\pi-\right]\end{pmatrix}\otimes\Omega_+\right]\begin{pmatrix}\Gamma_-^{-1}e_1 & 0 \\ 0 & \Gamma_-^{-1}e_1\end{pmatrix} \\
= & \frac{e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1}{f_x(\pi)}\begin{pmatrix} E\left[R^2|x=\pi+\right]+E\left[R^2|x=\pi-\right] & E\left[R\eta|x=\pi+\right]+E\left[R\eta|x=\pi-\right] \\ E\left[R\eta|x=\pi+\right]+E\left[R\eta|x=\pi-\right] & E\left[\eta^2|x=\pi+\right]+E\left[\eta^2|x=\pi-\right]\end{pmatrix}.
\end{aligned}$$

At last, by Slutsky' theorem,

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \overline{\alpha}_\pi(\pi) - B_n^f\right) \xrightarrow{d} N\left(0, \Sigma_f\right).$$

Case (ii). For $\mathfrak{p} = 0, \cdots, p$,

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} d_i(\pi) k_h\left(x_i - \pi\right) \overline{y}_i^*\right] \\
= & \int_0^\infty \int_{u/h}^M k(v) v^{\mathfrak{p}} \beta_\pi(\pi + vh - u) \alpha_\pi(\pi + vh - u) f_{x^*}(\pi + vh - u) f_u(u) dv du \\
& + \int_{-\infty}^0 \int_0^M k(v) v^{\mathfrak{p}} \beta_\pi(\pi + vh - u) \alpha_\pi(\pi + vh - u) f_{x^*}(\pi + vh - u) f_u(u) dv du \\
& + \int \int_0^M k(v) v^{\mathfrak{p}} \left[m_\pi(\pi + vh - u) + p_\pi(\pi + vh - u)\alpha_\pi(\pi + vh - u)\right] f_{x^*}(\pi + vh - u) f_u(u) dv du \\
\to & \gamma_{\mathfrak{p}}^+ f_x(\pi) \int_{-\infty}^0 \beta_\pi(\pi - u)\alpha_\pi(\pi - u) f_{u|x}(u|\pi) du \\
& + \gamma_{\mathfrak{p}}^+ f_x(\pi) \left\{E\left[p_\pi(\pi - u)\alpha_\pi(\pi - u)|x = \pi\right] + E\left[m_\pi(\pi - u)|x = \pi\right]\right\},
\end{aligned}$$

and

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} \left(1 - d_i(\pi)\right) k_h\left(x_i - \pi\right) \overline{y}_i^*\right] \\
\to & \gamma_{\mathfrak{p}}^- f_x(\pi) \int_{-\infty}^0 \beta_\pi(\pi - u)\alpha_\pi(\pi - u) f_{u|x}(u|\pi) dv du \\
& + \gamma_{\mathfrak{p}}^- f_x(\pi) \left\{E\left[p_\pi(\pi - u)\alpha_\pi(\pi - u)|x = \pi\right] + E\left[m_\pi(\pi - u)|x = \pi\right]\right\}.
\end{aligned}$$

So

$$\widehat{m}_+(\pi) - \widehat{m}_-(\pi) \xrightarrow{p} 0.$$

Similarly, $\widehat{p}_+(\pi) - \widehat{p}_-(\pi) \xrightarrow{p} 0$. In summary, the probability limit of $\widehat{\alpha}(\pi)$ is not well defined.

Case (iii). Define

$$B_{4n} = \frac{B_m^+ - B_m^-}{B_p^+ - B_p^-} - \alpha_\pi(\pi), \tag{21}$$

It can be shown that for $\mathfrak{p} = 0, \cdots, p$,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} d_i(\pi) k_h\left(x_i - \pi\right) \overline{y}_i & \xrightarrow{p} \gamma_{\mathfrak{p}}^+ f_{x^*}(\pi) \left\{\left[p_\pi(\pi) + \beta_\pi(\pi)\right]\alpha_\pi(\pi) + m_\pi(\pi)\right\}, \\
\frac{1}{n}\sum_{i=1}^n \left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} \left(1 - d_i(\pi)\right) k_h\left(x_i - \pi\right) \overline{y}_i & \xrightarrow{p} \gamma_{\mathfrak{p}}^- f_{x^*}(\pi) \left[p_\pi(\pi)\alpha_\pi(\pi) + m_\pi(\pi)\right],
\end{aligned}$$

so $B_m^+ - B_m^- \xrightarrow{p} \beta_\pi(\pi)\alpha_\pi(\pi)$. Similarly, $B_p^+ - B_p^- \xrightarrow{p} \beta_\pi(\pi)$. So $B_{4n} = o_p(1)$.

As to the asymptotic variance of $\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \overline{\alpha}_\pi(\pi) - B_{4n}\right)$, we can repeat the proof of Case (i) to show that Σ_4 is as specified in the theorem.

Case (iv). Define

$$B_{5n} = \frac{B_m^{*+} - B_m^{*-}}{B_p^{*+} - B_p^{*-}} - \alpha_\pi(\pi), \tag{22}$$

then

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi)-\alpha_\pi(\pi)-B_{5n}\right)=\frac{\left(B_p^{*+}-B_p^{*-}\right)\sqrt{nh}\left(V_m^{*+}-V_m^{*-}\right)-\left(B_m^{*+}-B_m^{*-}\right)\sqrt{nh}\left(V_p^{*+}-V_p^{*-}\right)}{\left(B_p^{*+}-B_p^{*-}\right)\left(B_p^{*+}-B_p^{*-}+V_p^{*+}-V_p^{*-}\right)}.$$

For $\mathfrak{p}=0,\cdots,p$,

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}} d_i(\pi)k_h\left(x_i-\pi\right)\overline{y}_i^*\right] \\
= & \int_0^\infty\int_{\frac{n^{-\delta}}{h}\epsilon}^M k(v)v^{\mathfrak{p}}\beta_\pi(\pi+vh-n^{-\delta}\epsilon)\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
& +\int_{-\infty}^0\int_0^M k(v)v^{\mathfrak{p}}\beta_\pi(\pi+vh-n^{-\delta}\epsilon)\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
& +\int\int_0^M k(v)v^{\mathfrak{p}}\left[m_\pi(\pi+vh-n^{-\delta}\epsilon)+p_\pi(\pi+vh-n^{-\delta}\epsilon)\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)\right]f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
\to & f_{x^*}(\pi)\begin{cases}\gamma_{\mathfrak{p}}^+\left[\beta_\pi(\pi)\alpha_\pi(\pi)+m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \overline{\gamma}_{\mathfrak{p}}^+\left(C\right)\beta_\pi(\pi)\alpha_\pi(\pi)+\gamma_{\mathfrak{p}}^+F_\epsilon(0)\beta_\pi(\pi)\alpha_\pi(\pi)+\gamma_{\mathfrak{p}}^+\left[m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \gamma_{\mathfrak{p}}^+F_\epsilon(0)\beta_\pi(\pi)\alpha_\pi(\pi)+\gamma_{\mathfrak{p}}^+\left[m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to\infty.\end{cases}
\end{aligned}$$

and

$$\begin{aligned}
& E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}\left(1-d_i(\pi)\right)k_h\left(x_i-\pi\right)\overline{y}_i^*\right] \\
= & \int_{-\infty}^0\int_{\frac{n^{-\delta}}{h}\epsilon}^0 k(v)v^{\mathfrak{p}}\beta_\pi(\pi+vh-n^{-\delta}\epsilon)\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
& +\int\int_{-M}^0 k(v)v^{\mathfrak{p}}\left[m_\pi(\pi+vh-n^{-\delta}\epsilon)+p_\pi(\pi+vh-n^{-\delta}\epsilon)\alpha_\pi(\pi+vh-n^{-\delta}\epsilon)\right]f_{x^*}(\pi+vh-n^{-\delta}\epsilon)f_\epsilon(\epsilon)dvd\epsilon \\
\to & f_{x^*}(\pi)\begin{cases}\gamma_{\mathfrak{p}}^-\left[m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \left[\gamma_{\mathfrak{p}}^-\left(C\right)F_\epsilon(0)-\overline{\gamma}_{\mathfrak{p}}^-\left(C\right)\right]\beta_\pi(\pi)\alpha_\pi(\pi)+\gamma_{\mathfrak{p}}^-\left[m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ \gamma_{\mathfrak{p}}^-F_\epsilon(0)\beta_\pi(\pi)\alpha_\pi(\pi)+\gamma_{\mathfrak{p}}^-\left[m_\pi(\pi)+p_\pi(\pi)\alpha_\pi(\pi)\right], & \text{if } \frac{n^{-\delta}}{h}\to\infty.\end{cases}
\end{aligned}$$

So

$$B_m^{*+}-B_m^{*-}\xrightarrow{p}\begin{cases}\beta_\pi(\pi)\alpha_\pi(\pi), & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \beta_\pi(\pi)\alpha_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right], & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ 0, & \text{if } \frac{n^{-\delta}}{h}\to\infty.\end{cases}$$

Similarly,

$$B_p^{*+}-B_p^{*-}\xrightarrow{p}\begin{cases}\beta_\pi(\pi), & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \beta_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right], & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty),\\ 0, & \text{if } \frac{n^{-\delta}}{h}\to\infty.\end{cases}$$

In summary,

$$B_{5n}\xrightarrow{p}\begin{cases}\frac{\beta_\pi(\pi)\alpha_\pi(\pi)}{\beta_\pi(\pi)}-\alpha_\pi(\pi)=0, & \text{if } \frac{n^{-\delta}}{h}\to 0,\\ \frac{\beta_\pi(\pi)\alpha_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right]}{\beta_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,p}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,p}^-(C)\right]}-\alpha_\pi(\pi)=0, & \text{if } \frac{n^{-\delta}}{h}\to C\in(0,\infty).\end{cases}$$

But when $\frac{n^{-\delta}}{h} \to \infty$, the probability limit of B_{5n} is not well defined.

As in Case 1, we need to derive the asymptotic variance of $\left(\sqrt{nh}\left(V_m^{*+} - V_m^{*-}\right), \sqrt{nh}\left(V_p^{*+} - V_p^{*-}\right)\right)$ to derive the asymptotic distribution of $\sqrt{nh}\left(\widehat{\alpha}_f(\pi) - \alpha_\pi(\pi) - B_{5n}\right)$. Now

$$\begin{pmatrix} \sqrt{nh}\left(V_m^+ - V_m^-\right) \\ \sqrt{nh}\left(V_p^+ - V_p^-\right) \end{pmatrix} = \begin{pmatrix} e_1' S_{n+}^{-1}(\pi) & 0 \\ 0 & e_1' S_{n+}^{-1}(\pi) \end{pmatrix} A_+^* - \begin{pmatrix} e_1' S_{n-}^{-1}(\pi) & 0 \\ 0 & e_1' S_{n-}^{-1}(\pi) \end{pmatrix} A_-^*.$$

where

$$A_+^* = \begin{pmatrix} \frac{1}{\sqrt{nh}} \sum_{i=1}^n Z_i(\pi) d_i(\pi) k\left(\frac{x_i-\pi}{h}\right) R_i^* \\ \frac{1}{\sqrt{nh}} \sum_{i=1}^n Z_i(\pi) d_i(\pi) k\left(\frac{x_i-\pi}{h}\right) \eta_i \end{pmatrix}, A_-^* = \begin{pmatrix} \frac{1}{\sqrt{nh}} \sum_{i=1}^n Z_i(\pi)\left(1-d_i(\pi)\right) k\left(\frac{x_i-\pi}{h}\right) R_i^* \\ \frac{1}{\sqrt{nh}} \sum_{i=1}^n Z_i(\pi)\left(1-d_i(\pi)\right) k\left(\frac{x_i-\pi}{h}\right) \eta_i \end{pmatrix}.$$

For $l = 0, \cdots, 2p$,

$$\begin{aligned}
& \frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi) R_i^{*2}\right] \\
= & \frac{1}{h} E\left[k^2\left(\frac{x^* + n^{-\delta}\epsilon - \pi}{h}\right)\left(\frac{x^* + n^{-\delta}\epsilon - \pi}{h}\right)^l 1\left(x^* + n^{-\delta}\epsilon \geq \pi\right) E\left[R^{*2}|x^*\right]\right] \\
= & \int\int_{-\frac{n^{-\delta}}{h}\epsilon}^M k^2\left(v + \frac{n^{-\delta}}{h}\epsilon\right)\left(v + \frac{n^{-\delta}}{h}\epsilon\right)^l E\left[R^{*2}|x^* = \pi + vh\right] f_{x^*}(\pi + vh) f_\epsilon(\epsilon) dv d\epsilon \\
\to & f_{x^*}(\pi) \int_0^\infty \int_{-C\epsilon}^0 k^2(v + C\epsilon)(v + C\epsilon)^l E\left[R^{*2}|x^* = \pi-\right] f_\epsilon(\epsilon) dv d\epsilon \\
& + f_{x^*}(\pi) \int_0^\infty \int_0^M k^2(v + C\epsilon)(v + C\epsilon)^l E\left[R^{*2}|x^* = \pi+\right] f_\epsilon(\epsilon) dv d\epsilon \\
& + f_{x^*}(\pi) \int_{-\infty}^0 \int_{-C\epsilon}^M k^2(v + C\epsilon)(v + C\epsilon)^l E\left[R^{*2}|x^* = \pi+\right] f_\epsilon(\epsilon) dv d\epsilon \\
= & f_{x^*}(\pi)\left[E\left[R^{*2}|x^* = \pi-\right] \overline{\overline{\omega}}_l^{+-}(C) + E\left[R^{*2}|x^* = \pi+\right] \overline{\overline{\omega}}_l^{++}(C)\right].
\end{aligned}$$

and similarly,

$$\begin{aligned}
\frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi) \eta_i^2\right] &\to f_{x^*}(\pi) E\left[\eta^2|x^* = \pi+\right] \omega_l^+, \\
\frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l d_i(\pi) R_i^* \eta_i\right] &\to f_{x^*}(\pi)\left[E\left[R^*\eta|x^* = \pi-\right] \overline{\overline{\omega}}_l^{+-}(C) + E\left[R^*\eta|x^* = \pi+\right] \overline{\overline{\omega}}_l^{++}(C)\right], \\
\frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l \left(1-d_i(\pi)\right) R_i^{*2}\right] &\to f_{x^*}(\pi)\left[E\left[R^{*2}|x^* = \pi-\right] \overline{\overline{\omega}}_l^{--}(C) + E\left[R^{*2}|x^* = \pi+\right] \overline{\overline{\omega}}_l^{-+}(C)\right], \\
\frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l \left(1-d_i(\pi)\right) \eta_i^2\right] &\to f_{x^*}(\pi) E\left[\eta^2|x^* = \pi-\right] \omega_l^-, \\
\frac{1}{h} E\left[k^2\left(\frac{x_i-\pi}{h}\right)\left(\frac{x_i-\pi}{h}\right)^l \left(1-d_i(\pi)\right) R_i^* \eta_i\right] &\to f_{x^*}(\pi)\left[E\left[R^*\eta|x^* = \pi-\right] \overline{\overline{\omega}}_l^{--}(C) + E\left[R^*\eta|x^* = \pi+\right] \overline{\overline{\omega}}_l^{-+}(C)\right].
\end{aligned}$$

As a result,

$$\begin{pmatrix} \sqrt{nh}\left(V_m^{*+}-V_m^{*-}\right) \\ \sqrt{nh}\left(V_p^{*+}-V_p^{*-}\right) \end{pmatrix} \xrightarrow{d} N\left(0,\Sigma_{mp}^{*}\right),$$

where

$$\begin{aligned} & f_{x^*}(\pi)\,\Sigma_{mp} \\ = & \begin{pmatrix} E\left[R^{*2}|x^*=\pi-\right]\Sigma_{+-}(C)+E\left[R^{*2}|x^*=\pi+\right]\Sigma_{++}(C) & E\left[R^{*}\eta|x^*=\pi-\right]\Sigma_{+-}(C)+E\left[R^{*}\eta|x^*=\pi+\right]\Sigma_{++}(C) \\ E\left[R^{*}\eta|x^*=\pi-\right]\Sigma_{+-}(C)+E\left[R^{*}\eta|x^*=\pi+\right]\Sigma_{++}(C) & E\left[\eta^2|x^*=\pi+\right]\Sigma \end{pmatrix} \\ & +\begin{pmatrix} E\left[R^{*2}|x^*=\pi-\right]\Sigma_{--}(C)+E\left[R^{*2}|x^*=\pi+\right]\Sigma_{-+}(C) & E\left[R^{*}\eta|x^*=\pi-\right]\Sigma_{--}(C)+E\left[R^{*}\eta|x^*=\pi+\right]\Sigma_{-+}(C) \\ E\left[R^{*}\eta|x^*=\pi-\right]\Sigma_{--}(C)+E\left[R^{*}\eta|x^*=\pi+\right]\Sigma_{-+}(C) & E\left[\eta^2|x^*=\pi-\right]\Sigma \end{pmatrix} \end{aligned}$$

At last, by Slutsky's theorem

$$\sqrt{nh}\left(\widehat{\alpha}_f(\pi)-\alpha_\pi(\pi)-B_{5n}\right) \xrightarrow{d} N\left(0,\Sigma_5(C)\right).$$

■

Proof of Theorem 8. $\overline{y}_i=\beta_\pi(x_i)\alpha_\pi(x_i-u_i)d_i(\pi)+m_\pi(x_i-u_i)+p_\pi(x_i)\alpha_\pi(x_i-u_i)$, and $\overline{D}_i=p_\pi(x_i)+\beta_\pi(x_i)d_i(\pi)$. To simplify notations, define $\mathbf{m}(\cdot)=m_\pi(\cdot)+\alpha_\pi(\cdot)$ and $\mathbf{p}(\cdot)=p_\pi(\cdot)+\beta_\pi(\cdot)$. Under Assumption CD, CY and U,

$$\begin{aligned} \beta_\pi(x_i)\alpha_\pi(x_i-u_i) &\approx \beta_\pi(\pi)\alpha_\pi(\pi)+\beta_\pi'(\pi)\alpha_\pi(\pi)(x_i-\pi)+\beta_\pi(\pi)\alpha_\pi'(\pi)(x_i-\pi-u_i) \\ &\quad+\frac{1}{2}\beta_\pi''(\pi)\alpha_\pi(\pi)(x_i-\pi)^2+\frac{1}{2}\beta_\pi(\pi)\alpha_\pi''(\pi)(x_i-\pi-u_i)^2+\beta_\pi'(\pi)\alpha_\pi'(\pi)(x_i-\pi)(x_i-\pi-u_i), \\ m_\pi(x_i-u_i) &\approx m_\pi(\pi)+m_\pi'(\pi)(x_i-\pi-u_i)+\frac{m_\pi''(\pi)}{2}(x_i-\pi-u_i)^2, \\ p_\pi(x_i)\alpha_\pi(x_i-u_i) &\approx p_\pi(\pi)\alpha_\pi(\pi)+p_\pi'(\pi)\alpha_\pi(\pi)(x_i-\pi)+p_\pi(\pi)\alpha_\pi'(\pi)(x_i-\pi-u_i) \\ &\quad+\frac{1}{2}p_\pi''(\pi)\alpha_\pi(\pi)(x_i-\pi)^2+\frac{1}{2}p_\pi(\pi)\alpha_\pi''(\pi)(x_i-\pi-u_i)^2+p_\pi'(\pi)\alpha_\pi'(\pi)(x_i-\pi)(x_i-\pi-u_i), \\ p_\pi(x_i) &\approx p_\pi(\pi)+p_\pi'(\pi)(x_i-\pi)+\frac{p_\pi''(\pi)}{2}(x_i-\pi)^2, \\ \beta_\pi(x_i) &\approx \beta_\pi(\pi)+\beta_\pi'(\pi)(x_i-\pi)+\frac{\beta_\pi''(\pi)}{2}(x_i-\pi)^2. \end{aligned}$$

By the discrete orthogonality relation,

$$\begin{aligned} B_m^+-B_m^- \approx\ & \beta_\pi(\pi)\alpha_\pi(\pi)+e_1'S_{n+}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n}Z_i(\pi)\,d_i(\pi)k_h(x_i-\pi)\left[-\left(\mathbf{p}(\pi)\alpha_\pi'(\pi)+m_\pi'(\pi)\right)u_i\right. \\ & \left.+\frac{1}{2}\mathbf{p}''(\pi)\alpha_\pi(\pi)(x_i-\pi)^2+\frac{\mathbf{p}(\pi)\alpha_\pi''(\pi)+m_\pi''(\pi)}{2}(x_i-\pi-u_i)^2+\mathbf{p}'(\pi)\alpha_\pi'(\pi)(x_i-\pi)(x_i-\pi-u_i)\right] \\ & -e_1'S_{n-}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n}Z_i(\pi)\left(1-d_i(\pi)\right)k_h(x_i-\pi)\left[-\left(p_\pi(\pi)\alpha_\pi'(\pi)+m_\pi'(\pi)\right)u_i\right. \\ & \left.+\frac{1}{2}p_\pi''(\pi)\alpha_\pi(\pi)(x_i-\pi)^2+\frac{p_\pi(\pi)\alpha_\pi''(\pi)+m_\pi''(\pi)}{2}(x_i-\pi-u_i)^2+p_\pi'(\pi)\alpha_\pi'(\pi)(x_i-\pi)(x_i-\pi-u_i)\right], \end{aligned}$$

and

$$B_p^+ - B_p^- \approx \beta_\pi(\pi) + e_1' S_{n+}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)\, d_i(\pi) k_h(x_i-\pi)\frac{\mathbf{p}''(\pi)}{2}(x_i-\pi)^2$$
$$-e_1' S_{n-}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^{n} Z_i(\pi)(1-d_i(\pi))\, k_h(x_i-\pi)\frac{p_\pi''(\pi)}{2}(x_i-\pi)^2.$$

From the analysis in the proof of Theorem 6, we have

$$\begin{aligned}
B_m^+ - B_m^- \approx\ & \beta_\pi(\pi)\alpha_\pi(\pi) + e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_+^{-1} - h\frac{f_{x^*}'(\pi)}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1} + n^{-\delta}\frac{f_{x^*}'(\pi)E[\epsilon]}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\right) \\
& \left[\begin{array}{c} -(\mathbf{p}(\pi)\alpha_\pi'(\pi)+m_\pi'(\pi))\left[f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^+E[\epsilon] + f_{x^*}'(\pi)\left(hn^{-\delta}\mu_{1,2}^+E[\epsilon] - n^{-2\delta}\mu_{0,1}^+E[\epsilon^2]\right)\right] \\ +\frac{1}{2}\mathbf{p}''(\pi)\alpha_\pi(\pi)f_{x^*}(\pi)h^2\mu_{2,3}^+ + \frac{\mathbf{p}(\pi)\alpha_\pi''(\pi)+m_\pi''(\pi)}{2}f_{x^*}(\pi)\left(h^2\mu_{2,3}^+ - 2hn^{-\delta}\mu_{1,2}^+E[\epsilon] + n^{-2\delta}\mu_{0,1}^+E[\epsilon^2]\right) \\ +\mathbf{p}'(\pi)\alpha_\pi'(\pi)f_{x^*}(\pi)\left(h^2\mu_{2,3}^+ - hn^{-\delta}\mu_{1,2}^+E[\epsilon]\right) \end{array}\right] \\
& -e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_-^{-1} - h\frac{f_{x^*}'(\pi)}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1} + n^{-\delta}\frac{f_{x^*}'(\pi)E[\epsilon]}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\right) \\
& \left[\begin{array}{c} -(p_\pi(\pi)\alpha_\pi'(\pi)+m_\pi'(\pi))\left[f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^-E[\epsilon] + f_{x^*}'(\pi)\left(hn^{-\delta}\mu_{1,2}^-E[\epsilon] - n^{-2\delta}\mu_{0,1}^-E[\epsilon^2]\right)\right] \\ +\frac{1}{2}p_\pi''(\pi)\alpha_\pi(\pi)f_{x^*}(\pi)h^2\mu_{2,3}^- + \frac{p_\pi(\pi)\alpha_\pi''(\pi)+m_\pi''(\pi)}{2}f_{x^*}(\pi)\left(h^2\mu_{2,3}^- - 2hn^{-\delta}\mu_{1,2}^-E[\epsilon] + n^{-2\delta}\mu_{0,1}^-E[\epsilon^2]\right) \\ +p_\pi'(\pi)\alpha_\pi'(\pi)f_{x^*}(\pi)\left(h^2\mu_{2,3}^- - hn^{-\delta}\mu_{1,2}^-E[\epsilon]\right) \end{array}\right] \\
=\ & \beta_\pi(\pi)\alpha_\pi(\pi) - n^{-\delta}\beta_\pi(\pi)\alpha_\pi'(\pi)E[\epsilon] + n^{-2\delta}\frac{f_{x^*}'(\pi)}{f_{x^*}(\pi)}\beta_\pi(\pi)\alpha_\pi'(\pi)E[\epsilon^2] + n^{-2\delta}\frac{\beta_\pi(\pi)\alpha_\pi''(\pi)}{2}E[\epsilon^2] \\
& -n^{-2\delta}\frac{f_{x^*}'(\pi)}{f_{x^*}(\pi)}E[\epsilon]^2\beta_\pi(\pi)\alpha_\pi'(\pi) + h^2\frac{\beta_\pi''(\pi)\alpha_\pi(\pi)+\beta_\pi(\pi)\alpha_\pi''(\pi)+2\beta_\pi'(\pi)\alpha_\pi'(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+ \\
=\ & \beta_\pi(\pi)\alpha_\pi(\pi) - n^{-\delta}\beta_\pi(\pi)\alpha_\pi'(\pi)E[\epsilon] + n^{-2\delta}\beta_\pi(\pi)\left[\frac{f_{x^*}'(\pi)}{f_{x^*}(\pi)}\alpha_\pi'(\pi)Var(\epsilon) + \frac{\alpha_\pi''(\pi)}{2}E[\epsilon^2]\right] \\
& +h^2\frac{\beta_\pi''(\pi)\alpha_\pi(\pi)+\beta_\pi(\pi)\alpha_\pi''(\pi)+2\beta_\pi'(\pi)\alpha_\pi'(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+,
\end{aligned}$$

and

$$B_p^+ - B_p^- \approx \beta_\pi(\pi) + h^2\frac{\beta_\pi''(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+.$$

So

$$\begin{aligned}
B_{4n} =\ & \frac{B_m^+ - B_m^-}{B_p^+ - B_p^-} - \alpha_\pi(\pi) \\
\approx\ & \frac{1}{\beta_\pi(\pi)}\left\{-n^{-\delta}\beta_\pi(\pi)\alpha_\pi'(\pi)E[\epsilon] + n^{-2\delta}\beta_\pi(\pi)\left[\frac{f_{x^*}'(\pi)}{f_{x^*}(\pi)}\alpha_\pi'(\pi)Var(\epsilon) + \frac{\alpha_\pi''(\pi)}{2}E[\epsilon^2]\right]\right. \\
& \left.+h^2\frac{\beta_\pi''(\pi)\alpha_\pi(\pi)+\beta_\pi(\pi)\alpha_\pi''(\pi)+2\beta_\pi'(\pi)\alpha_\pi'(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+\right\} - \frac{\alpha_\pi(\pi)\beta_\pi(\pi)}{\beta_\pi^2(\pi)}h^2\frac{\beta_\pi''(\pi)}{2}e_1'\Gamma_+^{-1}\mu_{2,3}^+,
\end{aligned}$$

which simplifies to the form stated in the theorem.

We now turn to the analysis of B_{5n}. $\overline{y}_i^* = \beta_\pi(x_i-u_i)\alpha_\pi(x_i-u_i)d_i^*(\pi)+m_\pi(x_i-u_i)+p_\pi(x_i-u_i)\alpha_\pi(x_i-u_i)$, and $\overline{D}_i^* \equiv p_\pi(x_i^*) + \beta_\pi(x_i^*)d_i^*(\pi)$. We only use the first order Taylor expansion in this case.

$$\begin{aligned}
\beta_\pi(x_i-u_i)\alpha_\pi(x_i-u_i) &\approx \beta_\pi(\pi)\alpha_\pi(\pi) + \left(\beta_\pi'(\pi)\alpha_\pi(\pi)+\beta_\pi(\pi)\alpha_\pi'(\pi)\right)(x_i-\pi-u_i), \\
p_\pi(x_i-u_i)\alpha_\pi(x_i-u_i) &\approx p_\pi(\pi)\alpha_\pi(\pi) + \left(p_\pi'(\pi)\alpha_\pi(\pi)+p_\pi(\pi)\alpha_\pi'(\pi)\right)(x_i-\pi-u_i), \\
m_\pi(x_i-u_i) &\approx m_\pi(\pi) + m_\pi'(\pi)(x_i-\pi-u_i).
\end{aligned}$$

Define

$$\begin{aligned}\Xi_1 &= \beta_\pi(\pi)\alpha_\pi(\pi),\ \Xi_1' = \beta_\pi'(\pi)\alpha_\pi(\pi) + \beta_\pi(\pi)\alpha_\pi'(\pi),\\ \Xi_0' &= m_\pi'(\pi) + p_\pi'(\pi)\alpha_\pi(\pi) + p_\pi(\pi)\alpha_\pi'(\pi),\end{aligned}$$

then

$$\begin{aligned}B_m^{*+} - B_m^{*-} \approx\ & e_1' S_{n+}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^n Z_i(\pi)\, d_i(\pi) k_h(x_i - \pi)\cdot\\ & \{d_i^*(\pi)[\Xi_1 + \Xi_1'(x_i - \pi - u_i)] - \Xi_0' u_i\}\\ & - e_1' S_{n-}^{-1}(\pi)\frac{1}{n}\sum_{i=1}^n Z_i(\pi)(1 - d_i(\pi))\, k_h(x_i - \pi)\cdot\\ & \{d_i^*(\pi)[\Xi_1 + \Xi_1'(x_i - \pi - u_i)] - \Xi_0' u_i\}.\end{aligned}$$

For $\mathfrak{p} = 0, 1$,

$$\begin{aligned}& E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} d_i(\pi) d_i^*(\pi) k_h(x_i - \pi)(x_i - \pi - u_i)\right]\\ =\ & \int_0^\infty\int_{\frac{n^{-\delta}}{h}\epsilon}^M k(v) v^{\mathfrak{p}}\left(vh - n^{-\delta}\epsilon\right) f_{x^*}(\pi + vh - n^{-\delta}\epsilon) f_\epsilon(\epsilon)\, dv\, d\epsilon\\ & + \int_{-\infty}^0\int_0^M k(v) v^{\mathfrak{p}}\left(vh - n^{-\delta}\epsilon\right) f_{x^*}(\pi + vh - n^{-\delta}\epsilon) f_\epsilon(\epsilon)\, dv\, d\epsilon\\ \approx\ & h f_{x^*}(\pi)\left[\overline{\gamma}_{\mathfrak{p}+1}^+(C) + \gamma_{\mathfrak{p}+1}^+ F_\epsilon(0)\right] - n^{-\delta} f_{x^*}(\pi)\left[\overline{\overline{\gamma}}_{\mathfrak{p}}^+(C) + \gamma_{\mathfrak{p}}^+\int_{-\infty}^0 \epsilon f_\epsilon(\epsilon)\, d\epsilon\right],\end{aligned}$$

and

$$\begin{aligned}& E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}} k_h(x_i - \pi)\, d_i(\pi) d_i^*(\pi)\right]\\ =\ & \int_0^\infty\int_{\frac{n^{-\delta}}{h}\epsilon}^M k(v) v^{\mathfrak{p}} f_{x^*}(\pi + vh - n^{-\delta}\epsilon) f_\epsilon(\epsilon)\, dv\, d\epsilon + \int_{-\infty}^0\int_0^M k(v) v^{\mathfrak{p}} f_{x^*}(\pi + vh - n^{-\delta}\epsilon) f_\epsilon(\epsilon)\, dv\, d\epsilon\\ =\ & f_{x^*}(\pi)\left(\overline{\gamma}_{\mathfrak{p}}^+(C) + \gamma_{\mathfrak{p}}^+ F_\epsilon(0)\right) + f_{x^*}'(\pi)\left[h\left(\overline{\gamma}_{\mathfrak{p}+1}^+(C) + \gamma_{\mathfrak{p}+1}^+ F_\epsilon(0)\right) - n^{-\delta}\left(\overline{\overline{\gamma}}_{\mathfrak{p}}^+(C) + \gamma_{\mathfrak{p}}^+\int_{-\infty}^0 \epsilon f_\epsilon(\epsilon)\, d\epsilon\right)\right].\end{aligned}$$

Similarly,

$$\begin{aligned}& E\left[\frac{1}{n}\sum_{i=1}^n\left(\frac{x_i - \pi}{h}\right)^{\mathfrak{p}}(1 - d_i(\pi))\, d_i^*(\pi)\, k_h(x_i - \pi)(x_i - \pi - u_i)\right]\\ \approx\ & f_{x^*}(\pi)\left[h\left(\gamma_{\mathfrak{p}+1}^- F_\epsilon(0) - \overline{\gamma}_{\mathfrak{p}+1}^-(C)\right) - n^{-\delta}\left(\gamma_{\mathfrak{p}}^-\int_{-\infty}^0 \epsilon f_\epsilon(\epsilon)\, d\epsilon - \overline{\overline{\gamma}}_{\mathfrak{p}}^+(C)\right)\right],\end{aligned}$$

and

$$E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\pi}{h}\right)^{\mathfrak{p}}k_h\left(x_i-\pi\right)\left(1-d_i(\pi)\right)d_i^*(\pi)\right]$$
$$\approx \quad f_{x^*}(\pi)\left[\gamma_{\mathfrak{p}}^- F_\epsilon(0)-\overline{\gamma}_{\mathfrak{p}}^-(C)\right]+f'_{x^*}(\pi)\left[h\left(\gamma_{\mathfrak{p}+1}^- F_\epsilon(0)-\overline{\gamma}_{\mathfrak{p}+1}^-(C)\right)-n^{-\delta}\left(\gamma_{\mathfrak{p}}^-\int_{-\infty}^{0}\epsilon f_\epsilon(\epsilon)d\epsilon-\overline{\overline{\gamma}}_{\mathfrak{p}}^-(C)\right)\right].$$

So

$$\begin{aligned}
B_m^{*+}-B_m^{*-} \quad \approx \quad & e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_+^{-1}-h\frac{f'_{x^*}(\pi)}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1}+n^{-\delta}\frac{f'_{x^*}(\pi)E\left[\epsilon\right]}{f_{x^*}^2(\pi)}\Gamma_+^{-1}\right)\\
& \left\{\Xi_1\left[f_{x^*}(\pi)\left(\overline{\mu}_{0,1}^+(C)+\mu_{0,1}^+F_\epsilon(0)\right)+f'_{x^*}(\pi)\Psi_R(h,n^{-\delta})\right]\right.\\
& \left.+\Xi_1' f_{x^*}(\pi)\Psi_R(h,n^{-\delta})-\Xi_0' f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^+E\left[\epsilon\right]\right\}\\
& -e_1'\left(\frac{1}{f_{x^*}(\pi)}\Gamma_-^{-1}-h\frac{f'_{x^*}(\pi)}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1}+n^{-\delta}\frac{f'_{x^*}(\pi)E\left[\epsilon\right]}{f_{x^*}^2(\pi)}\Gamma_-^{-1}\right)\\
& \left\{\Xi_1\left[f_{x^*}(\pi)\left(\mu_{0,1}^-F_\epsilon(0)-\overline{\mu}_{0,1}^-(C)\right)+f'_{x^*}(\pi)\Psi_L(h,n^{-\delta})\right]\right.\\
& \left.+\Xi_1' f_{x^*}(\pi)\Psi_L(h,n^{-\delta})-\Xi_0' f_{x^*}(\pi)n^{-\delta}\mu_{0,1}^-E\left[\epsilon\right]\right\},
\end{aligned}$$

where

$$\begin{aligned}
\Psi_R(h,n^{-\delta}) \quad &= \quad h\left(\overline{\mu}_{1,2}^+(C)+\mu_{1,2}^+F_\epsilon(0)\right)-n^{-\delta}\left(\overline{\overline{\mu}}_{0,1}^+(C)+\mu_{0,1}^+\int_{-\infty}^{0}\epsilon f_\epsilon(\epsilon)d\epsilon\right),\\
\Psi_L(h,n^{-\delta}) \quad &= \quad h\left(\mu_{1,2}^-F_\epsilon(0)-\overline{\mu}_{1,2}^+(C)\right)-n^{-\delta}\left(\mu_{0,1}^-\int_{-\infty}^{0}\epsilon f_\epsilon(\epsilon)d\epsilon-\overline{\overline{\mu}}_{0,1}^-(C)\right).
\end{aligned}$$

$B_m^{*+}-B_m^{*-}$ can be further simplified:

$$\begin{aligned}
B_m^{*+}-B_m^{*-} \quad \approx \quad & \Xi_1\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right]\\
& +\left(\Xi_1\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}+\Xi_1'\right)\left(e_1'\Gamma_+^{-1}\Psi_R(h,n^{-\delta})-e_1'\Gamma_-^{-1}\Psi_L(h,n^{-\delta})\right)\\
& +\Xi_1\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\left[\begin{array}{c}-h\left(e_1'\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right)\\+n^{-\delta}E\left[\epsilon\right]\left(e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right)\end{array}\right]\\
= \quad & \Xi_1\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right]\\
& +h\Xi_1'\left(e_1'\Gamma_+^{-1}\overline{\mu}_{1,2}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{1,2}^-(C)\right)-n^{-\delta}\Xi_1'\left(e_1'\Gamma_+^{-1}\overline{\overline{\mu}}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\overline{\mu}}_{0,1}^-(C)\right)\\
& +h\Xi_1\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\left[\left(e_1'\Gamma_+^{-1}\overline{\mu}_{1,2}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{1,2}^-(C)\right)-\left(e_1'\Gamma_+^{-1}\Gamma_{(+1)}\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\Gamma_{(-1)}\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right)\right]\\
& -n^{-\delta}\Xi_1\frac{f'_{x^*}(\pi)}{f_{x^*}(\pi)}\left[\left(e_1'\Gamma_+^{-1}\overline{\overline{\mu}}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\overline{\mu}}_{0,1}^-(C)\right)-E\left[\epsilon\right]\left(e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right)\right]\\
\equiv \quad & \Xi_1\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right]+\Psi(h,n^{-\delta};\Xi_1,\Xi_1').
\end{aligned}$$

$B_p^{*+}-B_p^{*-}$ can be approximated by a similar form except that Ξ_1 and Ξ_1' are replaced by $\beta_\pi(\pi)$ and $\beta'_\pi(\pi)$, respectively. Finally,

$$B_{5n}=\frac{B_m^{*+}-B_m^{*-}}{B_p^{*+}-B_p^{*-}}-\alpha_\pi(\pi)\approx\frac{\Psi(h,n^{-\delta};\Xi_1,\Xi_1')-\alpha_\pi(\pi)\,\Psi(h,n^{-\delta};\beta_\pi(\pi),\beta'_\pi(\pi))}{\beta_\pi(\pi)\left[e_1'\Gamma_+^{-1}\overline{\mu}_{0,1}^+(C)+e_1'\Gamma_-^{-1}\overline{\mu}_{0,1}^-(C)\right]}$$

which simplifies to the form stated in the theorem. ■

Supplementary Materials

1. Calculation in Example 2

In Example 2, $x^* \sim U[0,3]$. If $u \sim U[-1,1]$, then

$$\begin{aligned} f_{x,u}(x,u) &= \frac{1}{3}1(0 \leq x-u \leq 3)\frac{1}{2}1(-1 \leq u \leq 1) \\ &= \begin{cases} \frac{1}{6}1(-1 \leq u \leq x) & \text{if } x \leq 1, \\ \frac{1}{6}1(-1 \leq u \leq 1) & \text{if } 1 < x < 2, \\ \frac{1}{6}1(x-3 \leq u \leq 1) & \text{if } x > 2, \end{cases} \end{aligned}$$

and

$$f_x(x) = \int f_{x,u}(x,u)du = \begin{cases} \frac{x+1}{6}, & \text{if } x \leq 1, \\ \frac{2}{6}, & \text{if } 1 < x < 2, \\ \frac{4-x}{6}, & \text{if } x > 2. \end{cases}$$

So

$$E[u|x] = \frac{\int u f_{x,u}(x,u)du}{f_x(x)} = \begin{cases} \frac{x-1}{2}, & x \leq 1, \\ 0, & 1 < x < 2, \\ \frac{x-2}{2}, & x > 2. \end{cases}$$

When $u \sim N(0,1)$,

$$\begin{aligned} f_{x,u}(x,u) &= \frac{1}{3}1(0 \leq x-u \leq 3)\phi(u) = \frac{1}{3}1(x-3 \leq u \leq x)\phi(u), \\ f_x(x) &= \int f_{x,u}(x,u)du = \frac{1}{3}(\Phi(x) - \Phi(x-3)), \end{aligned}$$

where $\phi(u)$ and $\Phi(u)$ are the pdf and cdf of the standard normal distribution, respectively. So

$$\begin{aligned} E[u|x] &= \frac{\int u f_{x,u}(x,u)du}{f_x(x)} = \frac{\int_{x-3}^{x} u\phi(u)du}{\Phi(x) - \Phi(x-3)} \\ &= \frac{\phi(x-3) - \phi(x)}{\Phi(x) - \Phi(x-3)}, \end{aligned}$$

where the last equality is from the fact that $\int_x^\infty u\phi(u)du = \phi(x)$.

2. Extension of Goldberger (2008)

Suppose the forcing variable is x^* in the sharp design. First, we find $E[y|d_\pi^* = 1, x = \pi+] - E[y|d_\pi^* = 0, x = \pi-]$. For this purpose, we need to develop the conditional densities $f_{x^*|d_\pi^*,x}(x^*|d_\pi^*, x)$ and $f_{\varepsilon|d_\pi^*,x}(\varepsilon|d_\pi^*, x)$

since $y = m_\pi^*(x^*) + \alpha_\pi d_\pi^* + \varepsilon$ is a function of x^* and ε. For $d_\pi^* = 1$,

$$\begin{aligned}
f_{x^*|d_\pi^*,x}\left(x^*|d_\pi^*=1,x\right) &= \frac{f_{x^*,x|d_\pi^*}\left(x^*,x|d_\pi^*=1\right)}{f_{x|d_\pi^*}\left(x|d_\pi^*=1\right)} = \frac{f_{x|x^*,d_\pi^*}\left(x|x^*,d_\pi^*=1\right)f_{x^*|d_\pi^*}\left(x^*|d_\pi^*=1\right)}{f_{x|d_\pi^*}\left(x|d_\pi^*=1\right)} \\
&= \frac{f_{x|x^*}\left(x|x^*\right)1\left(x^*\geq\pi\right)f_{x^*|d_\pi^*}\left(x^*|d_\pi^*=1\right)}{f_{x|d_\pi^*}\left(x|d_\pi^*=1\right)} \\
&= \frac{f_u(x-x^*)1\left(x^*\geq\pi\right)f_{x^*}\left(x^*\right)/(1-F_{x^*}(\pi))}{\int_\pi^\infty f_u(x-x^*)f_{x^*}\left(x^*\right)/(1-F_{x^*}(\pi))dx^*} \\
&= \frac{f_{x^*}\left(x^*\right)f_u(x-x^*)}{\int_\pi^\infty f_{x^*}\left(x^*\right)f_u(x-x^*)dx^*}1\left(x^*\geq\pi\right),
\end{aligned}$$

so

$$E\left[m_\pi^*(x^*)|d_\pi^*=1,x\right] = \frac{\int_\pi^\infty m_\pi^*(x^*)f_{x^*}\left(x^*\right)f_u(x-x^*)dx^*}{\int_\pi^\infty f_{x^*}\left(x^*\right)f_u(x-x^*)dx^*} = \frac{\int_\pi^\infty m_\pi^*(x^*)f_{x^*|x}\left(x^*|x\right)dx^*}{1-F_{x^*|x}\left(\pi|x\right)}.$$

Symmetrically,

$$E\left[m_\pi^*(x^*)|d_\pi^*=0,x\right] = \frac{\int_{-\infty}^\pi m_\pi^*(x^*)f_{x^*}\left(x^*\right)f_u(x-x^*)dx^*}{\int_{-\infty}^\pi f_{x^*}\left(x^*\right)f_u(x-x^*)dx^*} = \frac{\int_{-\infty}^\pi m_\pi^*(x^*)f_{x^*|x}\left(x^*|x\right)dx^*}{F_{x^*|x}\left(\pi|x\right)}.$$

We then develop the conditional density $f_{\varepsilon|d_\pi^*,x}\left(\varepsilon|d_\pi^*,x\right)$. For $d_\pi^*=1$, $\varepsilon=\varepsilon_1$, then

$$\begin{aligned}
f_{\varepsilon|d_\pi^*,x}\left(\varepsilon|d_\pi^*=1,x\right) &= \frac{f_{\varepsilon,x|d_\pi^*}\left(\varepsilon,x|d_\pi^*=1\right)}{f_{x|d_\pi^*}\left(x|d_\pi^*=1\right)} = \frac{\int_\pi^\infty f_{x^*,\varepsilon}^1\left(x^*,\varepsilon\right)f_u(x-x^*)dx^*/(1-F_{x^*}(\pi))}{\int_\pi^\infty f_{x^*}(x^*)f_u(x-x^*)dx^*/(1-F_{x^*}(\pi))} \\
&= \frac{\int_\pi^\infty f_{x^*,\varepsilon}^1\left(x^*,\varepsilon\right)f_u(x-x^*)dx^*}{\int_\pi^\infty f_{x^*}(x^*)f_u(x-x^*)dx^*},
\end{aligned}$$

so

$$\begin{aligned}
E[\varepsilon|d_\pi^* &= 1,x] = \frac{\int\int_\pi^\infty \varepsilon f_{x^*,\varepsilon}^1\left(x^*,\varepsilon\right)f_u(x-x^*)dx^*d\varepsilon}{\int_\pi^\infty f_{x^*}(x^*)f_u(x-x^*)dx^*} \\
&= \frac{\int_\pi^\infty E\left[\varepsilon_1|x^*\right]f_{x^*}(x^*)f_u(x-x^*)dx^*}{\int_\pi^\infty f_{x^*}(x^*)f_u(x-x^*)dx^*} = 0,
\end{aligned}$$

where the second equality is from Fubini's theorem, and the last equality is from $E\left[\varepsilon_1|x^*\right]=0$. Symmetrically, $E[\varepsilon|d_\pi^*=0,x]=0$.

In Example 3,

$$\begin{aligned}
\frac{\int_\pi^\infty m_\pi^*(x^*)f_u(x-x^*)f_{x^*}\left(x^*\right)dx^*}{\int_\pi^\infty f_u(x-x^*)f_{x^*}\left(x^*\right)dx^*} &= \frac{\int_\pi^\infty\left(\overline{\beta}_0+\overline{\beta}_1x^*\right)\phi(x-x^*;0,\frac{1-\rho}{\rho}\sigma^2)\phi(x^*;\mu,\sigma^2)dx^*}{\int_\pi^\infty\phi(x-x^*;0,\frac{1-\rho}{\rho}\sigma^2)\phi(x^*;\mu,\sigma^2)dx^*} \\
&= \overline{\beta}_0+\overline{\beta}_1\frac{\int_\pi^\infty x^*\phi(x-x^*;0,\frac{1-\rho}{\rho}\sigma^2)\phi(x^*;\mu,\sigma^2)dx^*}{\int_\pi^\infty\phi(x-x^*;0,\frac{1-\rho}{\rho}\sigma^2)\phi(x^*;\mu,\sigma^2)dx^*} \\
&= \overline{\beta}_0+\overline{\beta}_1\frac{\int_\pi^\infty x^*\phi\left(x^*;a(x),b^2\right)dx^*}{\int_\pi^\infty\phi\left(x^*;a(x),b^2\right)dx^*} \\
&= \overline{\beta}_0+\overline{\beta}_1\left[a(x)+b\lambda\left(\frac{a(x)-\pi}{b}\right)\right],
\end{aligned}$$

where $\phi\left(x;\mu,\sigma^2\right)$ is the density of a normal distribution with mean μ and variance σ^2, $\phi\left(x\right)=\phi\left(x;0,1\right)$, the third equality uses the facts that

$$\phi\left(x - x^*; 0, \frac{1-\rho}{\rho}\sigma^2\right)\phi(x^*; \mu, \sigma^2) = \phi\left(x^*; a(x), b^2\right)\phi\left(x; \mu, \frac{\sigma^2}{\rho}\right),$$

and the fourth inequality uses the fact that $E[X|X > \pi] = \mu + \sigma\lambda(c)$, where $X \sim N\left(\mu, \sigma^2\right)$, and $c = \frac{\mu-\pi}{\sigma}$. Similarly,

$$\frac{\int_{-\infty}^{\pi} m_\pi^*(x^*) f_u(\pi - x^*) f_{x^*}(x^*)\, dx^*}{\int_{-\infty}^{\pi} f_u(\pi - x^*) f_{x^*}(x^*)\, dx^*} = \beta_0 + \beta_1\left(a(x) - b\lambda\left(\frac{\pi - a(x)}{b}\right)\right).$$

Appendix C of Goldberger (2008) shows the special case with $\mu = \pi = \beta_0 = \alpha_0 = \alpha_1 = 0$ and $\beta_1 = 1$.

In this example, $u|x \sim N(x - a(x), b^2) \equiv N(\mu_{u|x}, b^2)$, so the propensity score

$$p(x) = F_{u|x}(x - \pi) = \Phi\left(\frac{x - \pi - \mu_{u|x}}{b}\right) = \Phi\left(\frac{a(x) - \pi}{b}\right),$$

and

$$\begin{aligned}
E[y|x] &= E\left[m_\pi^*(x^*) + \alpha_\pi d_\pi^*|x\right] = E\left[m_\pi^*(x^*)|x\right] + \alpha_\pi p(x) \\
&= \int_{-\infty}^{x-\pi}\left(\overline{\beta}_0 + \overline{\beta}_1(x-u)\right)\phi\left(u; \mu_{u|x}, b^2\right)du + \int_{x-\pi}^{\infty}\left(\beta_0 + \beta_1(x-u)\right)\phi\left(u; \mu_{u|x}, b^2\right)du \\
&\quad + \alpha_\pi \int_{-\infty}^{x-\pi}\phi\left(u; \mu_{u|x}, b^2\right)du \\
&= \left(\overline{\beta}_0 + \overline{\beta}_1 x\right)\Phi\left(\frac{x-\pi-\mu_{u|x}}{b}\right) - \overline{\beta}_1\left((x-a(x))\Phi\left(\frac{x-\pi-\mu_{u|x}}{b}\right) - b\phi\left(\frac{x-\pi-\mu_{u|x}}{b}\right)\right) \\
&\quad + \left(\beta_0 + \beta_1 x\right)\Phi\left(\frac{\mu_{u|x} - (x-\pi)}{b}\right) - \beta_1\left((x-a(x))\Phi\left(\frac{\mu_{u|x} - (x-\pi)}{b}\right) + b\phi\left(\frac{x-\pi-\mu_{u|x}}{b}\right)\right) \\
&\quad + \alpha_\pi \Phi\left(\frac{x-\pi-\mu_{u|x}}{b}\right) \\
&= \left(\overline{\beta}_0 + \overline{\beta}_1 a(x)\right)\Phi\left(\frac{a(x)-\pi}{b}\right) + \left(\beta_0 + \beta_1 a(x)\right)\left(1 - \Phi\left(\frac{a(x)-\pi}{b}\right)\right) \\
&\quad + \left(\overline{\beta}_1 - \beta_1\right)b\phi\left(\frac{a(x)-\pi}{b}\right) + \alpha_\pi b\Phi\left(\frac{a(x)-\pi}{b}\right) \\
&= \left[\beta_0 + \beta_1 a(x)\right] + \alpha_1\left(a(x) - \pi\right)\Phi\left(\frac{a(x)-\pi}{b}\right) + \alpha_1 b\phi\left(\frac{a(x)-\pi}{b}\right) + \alpha_\pi \Phi\left(\frac{a(x)-\pi}{b}\right),
\end{aligned}$$

which is a continuous function of x.

3. Simplification of the Asymptotic Variance in Case 5

This section can be treated as an example of many simplifications in the main text. Suppose $f_\epsilon(\cdot)$ is symmetric, then we can show $\overline{\overline{\omega}}_j^{+-}(C) = (-1)^j \overline{\overline{\omega}}_j^{-+}(C)$, and $\overline{\overline{\omega}}_j^{++}(C) = (-1)^j \overline{\overline{\omega}}_j^{--}(C)$. We only show

$\overline{\overline{\omega}}_j^{+-}(C) = (-1)^j \overline{\overline{\omega}}_j^{-+}(C)$ for illustration.

$$
\begin{aligned}
\overline{\overline{\omega}}_j^{+-}(C) &= \int_0^\infty \int_{-C\epsilon}^0 k^2(v+C\epsilon)(v+C\epsilon)^j f_\epsilon(\epsilon) dv d\epsilon \\
&= \int_{-\infty}^0 \int_{C\epsilon}^0 k^2(v-C\epsilon)(v-C\epsilon)^j f_\epsilon(-\epsilon) dv d\epsilon \\
&= \int_{-\infty}^0 \int_0^{-C\epsilon} k^2(-v-C\epsilon)(-v-C\epsilon)^j f_\epsilon(-\epsilon) dv d\epsilon \\
&= (-1)^j \int_{-\infty}^0 \int_0^{-C\epsilon} k^2(v+C\epsilon)(v+C\epsilon)^j f_\epsilon(\epsilon) dv d\epsilon = (-1)^j \overline{\overline{\omega}}_j^{--}(C).
\end{aligned}
$$

where the second and third equalities are from changing variables, and the second to last equality is from the symmetricity of $k(\cdot)$ and $f_\epsilon(\cdot)$. Given that

$$
\begin{aligned}
\Sigma_{+-}(C) &= e_1'\Gamma_+^{-1}\overline{\overline{\Omega}}_{+-}(C)\Gamma_+^{-1}e_1, \Sigma_{++}(C) = e_1'\Gamma_+^{-1}\overline{\overline{\Omega}}_{++}(C)\Gamma_+^{-1}e_1, \\
\Sigma_{-+}(C) &= e_1'\Gamma_-^{-1}\overline{\overline{\Omega}}_{-+}(C)\Gamma_-^{-1}e_1, \Sigma_{--}(C) = e_1'\Gamma_-^{-1}\overline{\overline{\Omega}}_{--}(C)\Gamma_-^{-1}e_1,
\end{aligned}
$$

and $\gamma_j^- = (-1)^j \gamma_j^+$, it is easy to show $\Sigma_{+-}(C) = \Sigma_{-+}(C)$, and $\Sigma_{++}(C) = \Sigma_{--}(C)$ by the same steps in showing $e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1 = e_1'\Gamma_-^{-1}\Omega_-\Gamma_-^{-1}e_1$.

When ε_0, ε_1, η_0 and η_1 are independent of each other conditional on x^*, we can simplify $E\left[R_0^{*2}|x^*=\pi\right]$ as follows:

$$
\begin{aligned}
E\left[R_0^{*2}|x^*=\pi\right] &= E\left[\eta_0^2|x^*=\pi\right]\left(\sigma_0^2(\pi)+\sigma_1^2(\pi)\right) + \alpha_\pi(\pi)^2 E\left[\eta_0^2|x^*=\pi\right] + p_\pi^2(\pi)\sigma_1^2(\pi) + (1-p_\pi(\pi))^2\sigma_0^2(\pi) \\
&= \alpha_\pi^2(\pi)E\left[\eta_0^2|x^*=\pi\right] + \left(p_\pi^2(\pi)+E\left[\eta_0^2|x^*=\pi\right]\right)\sigma_1^2(\pi) + \left[(1-p_\pi(\pi))^2 + E\left[\eta_0^2|x^*=\pi\right]\right]\sigma_0^2(\pi) \\
&= \alpha_\pi^2(\pi)E\left[\eta_0^2|x^*=\pi\right] + p_\pi(\pi)\sigma_1^2(\pi) + (1-p_\pi(\pi))\sigma_0^2(\pi)
\end{aligned}
$$

where the last equality is from $E\left[\eta_0^2|x^*=\pi\right] = p_\pi(\pi)(1-p_\pi(\pi))$. Similarly,

$$
\begin{aligned}
E\left[R_1^{*2}|x^*=\pi\right] &= \alpha_\pi^2(\pi)E\left[\eta_1^2|x^*=\pi\right] + (p_\pi(\pi)+\beta_\pi(\pi))\sigma_1^2(\pi) + (1-p_\pi(\pi)-\beta_\pi(\pi))\sigma_0^2(\pi) \\
E\left[R_0^*\eta_0|x^*=\pi\right] &= \alpha_\pi(\pi)E\left[\eta_0^2|x^*=\pi\right], E\left[R_1^*\eta_1|x^*=\pi\right] = \alpha_\pi(\pi)E\left[\eta_1^2|x^*=\pi\right].
\end{aligned}
$$

Now,

$$
\begin{aligned}
\Sigma_4 &= \frac{e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1}{f_{x^*}(\pi)\beta_\pi^2(\pi)}\left\{\left[E\left[R_0^{*2}|x^*=\pi\right]+E\left[R_1^{*2}|x^*=\pi\right]\right]\right. \\
&\quad -2\alpha_\pi(\pi)\left[E\left[R_0^*\eta_0|x^*=\pi\right]+E\left[R_1^*\eta_0|x^*=\pi\right]\right] \\
&\quad \left.+\alpha_\pi(\pi)^2\left[E\left[\eta_0^2|x^*=\pi\right]+E\left[\eta_1^2|x^*=\pi\right]\right]\right\} \\
&= \frac{e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1}{f_{x^*}(\pi)\beta_\pi^2(\pi)}\left\{\alpha_\pi^2(\pi)E\left[\eta_0^2|x^*=\pi\right] + p_\pi(\pi)\sigma_1^2(\pi) + (1-p_\pi(\pi))\sigma_0^2(\pi)\right. \\
&\quad +\alpha_\pi^2(\pi)E\left[\eta_1^2|x^*=\pi\right] + (p_\pi(\pi)+\beta_\pi(\pi))\sigma_1^2(\pi) + (1-p_\pi(\pi)-\beta_\pi(\pi))\sigma_0^2(\pi) \\
&\quad \left.-\alpha_\pi(\pi)^2\left[E\left[\eta_0^2|x^*=\pi\right]+E\left[\eta_1^2|x^*=\pi\right]\right]\right\} \\
&= \frac{e_1'\Gamma_+^{-1}\Omega_+\Gamma_+^{-1}e_1}{f_{x^*}(\pi)\beta_\pi^2(\pi)}\left[(2p_\pi(\pi)+\beta_\pi(\pi))\sigma_1^2(\pi) + (2-2p_\pi(\pi)-\beta_\pi(\pi))\sigma_0^2(\pi)\right].
\end{aligned}
$$

If $\sigma_0^2(\pi) = \sigma_1^2(\pi) = \sigma^2$,

$$\Sigma_4 = \frac{2\sigma^2}{f_{x^*}(\pi)\,\beta_\pi^2(\pi)} e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1.$$

$$\begin{aligned}
\Sigma_5(C) &= \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \left\{ E\left[R_0^{*2} | x^* = \pi\right] + E\left[R_1^{*2} | x^* = \pi\right] - \right. \\
&\quad -2\alpha_\pi(\pi)\left[E\left[R_0^* \eta_0 | x^* = \pi\right] + E\left[R_1^* \eta_1 | x^* = \pi\right]\right]\} \\
&\quad + \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1 \alpha_\pi(\pi)^2}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \left\{ E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right] \right\} \\
&= \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \left\{ \alpha_\pi^2(\pi) E\left[\eta_0^2 | x^* = \pi\right] + p_\pi(\pi)\sigma_1^2(\pi) + (1 - p_\pi(\pi))\,\sigma_0^2(\pi) \right. \\
&\quad + \alpha_\pi^2(\pi) E\left[\eta_1^2 | x^* = \pi\right] + (p_\pi(\pi) + \beta_\pi(\pi))\,\sigma_1^2(\pi) + (1 - p_\pi(\pi) - \beta_\pi(\pi))\,\sigma_0^2(\pi) \\
&\quad \left. -2\alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right] \right\} \\
&\quad + \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1 \alpha_\pi(\pi)^2}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right] \\
&= \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \left\{ (2p_\pi(\pi) + \beta_\pi(\pi))\,\sigma_1^2(\pi) + (2 - 2p_\pi(\pi) - \beta_\pi(\pi))\,\sigma_0^2(\pi) \right. \\
&\quad \left. -\alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right] \right\} \\
&\quad + \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1 \alpha_\pi(\pi)^2}{4 f_{x^*}(\pi)\,\beta_\pi^2(\pi) \left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right].
\end{aligned}$$

If $\sigma_0^2(\pi) = \sigma_1^2(\pi) = \sigma^2$,

$$\begin{aligned}
\Sigma_5(C) &= \frac{2\sigma^2 - \alpha_\pi^2(\pi)\left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right]}{f_{x^*}(\pi)\,\beta_\pi^2(\pi)} \frac{\Sigma_{+-}(C) + \Sigma_{++}(C)}{4\left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2} \\
&\quad + \frac{\alpha_\pi(\pi)^2 \left[E\left[\eta_0^2 | x^* = \pi\right] + E\left[\eta_1^2 | x^* = \pi\right]\right]}{f_{x^*}(\pi)\,\beta_\pi^2(\pi)} \frac{e_1' \Gamma_+^{-1} \Omega_+ \Gamma_+^{-1} e_1}{4\left(e_1' \Gamma_+^{-1} \overline{\mu}_{0,p}^{+}(C)\right)^2}.
\end{aligned}$$