UNIVERSITY OF CALIFORNIA SAN DIEGO

Econometric Analysis of Unconditional Policy Effects

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Julián Martínez-Iriarte

Committee in charge:

Professor Yixiao Sun, Chair
Professor Prashant Bharadwaj
Professor Jelena Bradic
Professor Xinwei Ma
Professor Augusto Nieto-Barthaburu
Professor Kaspar Wuthrich

2021

The Dissertation of Julián Martínez-Iriarte is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To my mother and my family. To the memory of my father.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my advisor Yixiao Sun for his constant support and generosity. It has been an honor to work under his guidance and learn econometrics from him. I continue to be inspired by his deep intellectual curiosity.

Kaspar Wuthrich took an early interest in my research and was always very generous with his time. His way of thinking has had a tremendous impact on every aspect of my research. Discussing ideas all over the wall of his office was always a fun and illuminating experience. Xinwei Ma joined my commitee the minute he arrived at UCSD, and has proven to be invaluable ever since. He provided extensive feedback on every chapter in this dissertation. The comments I received from Jelena Bradic and Prashant Bharadwaj have been extremely useful in the early stages of Chapters 1 and 2. Last, but not least, Augusto Nieto-Barthaburu provided very valuable insights on every chapter. Augusto's relentless optimism and moral support was very important, especially on years 4 through 6. Ever since I took his general equilibrium class in Tucumán, he always encouraged me to pursue a Ph.D.

I would like to thank my professors at UCSD, UTDT, and UNT. I would like to give special acknowledgement to Paula Juez Pérez, my high school professor, who inspired me to study economics in college. Ana María Cerro and Víctor J. Elías deserve credit for first exposing me to econometrics as an undergraduate student. During the PhD, I learned a lot from my UCSD classmates: Bruno, Ale, Camila, Daniel, Alex, Sameem, Lam, Pietro, and Pablo.

On a personal level, I want to thank my family, my friends, and my classmates. I am grateful for Meghan's love and unconditional support.

Chapters 1 and 2, in part, are being prepared for submission for publication, and are coauthored with Yixiao Sun. The dissertation author, Julián Martínez-Iriarte, was the primary author of these chapters.

Chapter 3 is currently being prepared for submission for publication of the material. The dissertation author, Julián Martínez-Iriarte, was the sole author of this material.

VITA

2011	B. S. in Economics, Universidad Nacional de Tucumán, Tucumán, Argentina
2014	Graduate Studies in Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina
2017	M. A. in Economics, University of California San Diego
2021	Ph. D. in Economics, University of California San Diego

PUBLICATIONS

Julián Martínez-Iriarte, Yixiao Sun, Xuexin Wang, "Asymptotic F Tests under Possibly Weak Identification", *Journal of Econometrics*, 218, 2020.

ABSTRACT OF THE DISSERTATION

Econometric Analysis of Unconditional Policy Effects

by

Julián Martínez-Iriarte

Doctor of Philosophy in Economics

University of California San Diego, 2021

Professor Yixiao Sun, Chair

This dissertation contributes to the analysis of the unconditional effects of counterfactual policies that manipulate the distribution of covariates. By unconditional effects we mean the effects on any functional of the unconditional distribution of the outcome.

Chapter 1 focuses on the effect on the unconditional quantiles of the outcome. We first show how to achieve identification under unconfoundedness. Then, we characterize the asymptotic bias of the unconditional regression estimator that ignores the endogeneity and elaborate on the channels that the endogeneity can render the unconditional regression estimator inconsistent. We show that even if the treatment status is exogenous, the unconditional regression estimator can still be inconsistent when there are common covariates affecting both the treatment

status and the outcome variable.

Chapter 2 provides identification and estimation results for the case of an endogenous binary variable. We introduce a new class of marginal treatment effects (MTE) based on the influence function of the functional underlying the policy target. We show that an unconditional policy effect can be represented as a weighted average of the newly defined MTEs over the individuals at the margin of indifference. Point identification is achieved using the local instrumental variable approach. Furthermore, the unconditional policy effects are shown to include the marginal policy-relevant treatment effect in the literature as a special case. Methods of estimation and inference for the unconditional policy effects are provided. In the empirical application, we estimate the effect of changing college enrollment status, induced by higher tuition subsidy, on the quantiles of the wage distribution.

Chapter 3 proposes a framework to analyze the effects of counterfactual policies when neither unconfoundedness holds nor an instrumental variable is available. For a given counterfactual policy, we obtain identified sets for the effect of both marginal and global changes in the proportion of treated individuals. To conduct a sensitivity analysis, we introduce the quantile breakdown frontier, a curve that quantifies the maximum amount of selection bias consistent with a given conclusion. To illustrate our method, we perform a sensitivity analysis on the effect of unionizing low-income workers on the quantiles of the distribution of wages.

Chapter 1

Bias in Unconditional Effects

1.1 Introduction

In this chapter we analyze different kinds of biases that arise in the estimation of unconditional causal effects as a result of the manipulation in the distribution of covariates. By unconditional effects we mean the effects on any functional of the unconditional distribution of the outcome. For concreteness, we focus on the effect on quantiles. Thus, we asses the effect on the unconditional quantiles of the outcome as a result of a counterfactual policy that alters the distribution of covariates. Consider a policy that increases unionization. The effect of such a policy on the median of wages (of all workers) is an example of an unconditional effect. We do not look at the median of each subpopulation separately; instead we pool all workers, unionized and nonunionized, and analyze the effect on the median of this pooled, unconditional distribution.

We work with a general model $Y = r(W,U)$, where W are observed covariates, and U are unobservables. The thought experiment is that of a policy maker that wants to manipulate a target variable in W and is interested on the impact of the manipulation in a function of the unconditional distribution[1] of Y. We distinguish two cases: when the target variables is binary, and when the target variable is continuous. In the former case, we write $W = (D,X)$, and D is

[1] This is also the marginal distribution. But we employ the word marginal to denote a certain type of effect to be defined later.

the binary target variable. When the target is continuous, we write $W = (X, \tilde{W})$, where X is the target variable, and $\tilde{W}$ are other observed covariates.

In both cases, interventions take the form of manipulations of the mean. However, the mechanism by which these manipulations operate differ markedly in each case. In the case of a continuous target variable, such manipulation takes the form of a shift. That is, we analyze a situation where X is replaced by $X + \delta$ for some quantity δ. For example, every individual in our sample receives (in a counterfactual fashion) extra years of education. In the case of a discrete target variable D, we cannot perform such shifts, since $D + \delta$ makes no sense. Here we try two approaches. The first one considers a setting where D is unconfounded and the counterfacual D, denoted D_δ, is unconfounded as well. The parameter δ reflects a mean shift: $\Pr(D_\delta = 1) = \Pr(D = 1) + \delta$. A second option is to rely in additional equation which explains the value of D as a function of another continuous random variable. It is this underlying continuous random variable which we intervene with a shift. For example, is D is college attendance, and Z is college tuition, then we can think of a counterfactual policy where Z is shifted to $Z + \delta$. In this case, δ can be interpreted as a tuition subsidy.

For both cases we will analyze situations where commonly used identifying assumptions break down. Moreover, we will provide closed form expression for the bias. This is of interest because when working with observational data, this is likely to happen. Section 1.2 analyzes the discrete case and Section 1.3 analyzes the continuous case. Both sections introduce novel results in terms of identification and closed form solutions for the bias. Finally Section 1.4 concludes. All proofs are relegated to the Appendix.

1.2 Discrete Target Variable

We will work with the potential outcomes framework. Let D be treatment status taking values 0 and 1. For some unknown functions r_0 and r_1

$$Y(0) = r_0(X, U_0),$$
$$Y(1) = r_1(X, U_1),$$

where X are observed covariates and U_0 and U_1 consist of unobservables. We do not impose any restriction on the dimension of the unobservables. As usual, we only observe either $Y = Y(1)$ for those individuals whose $D = 1$, or $Y = Y(0)$ for those individuals whose $D = 0$. Thus, the observed outcome can be written as

$$Y = D \cdot r_1(X, U_1) + (1 - D) \cdot r_0(X, U_0) := r(D, X, U), \tag{1.1}$$

for a general nonseparable function r, and $U := (U_0, U_1)'$. Hence, the potential outcomes framework can be cast into a structural modeling framework with a special causal function r. We will denote the proportion of treated individuals by $p := \Pr(D = 1)$. Here, $W = (D, X)'$.

In the rest of this dissertation, we maintain a continuity assumption about the outcome Y. This is not essential to our results, but allows us to reduce the notational burden. In particular, the assumption of a positive density is equivalent to uniqueness of the quantile, which is very convenient.

Assumption 1.1 (Continuity)**.** *The observed outcome Y is continuous, with positive density in its support $\mathscr{Y}$.*

A counterfactual policy is an alternative assignment of individuals to treatment. It is given by a binary random variable D_δ, such that $\Pr(D_\delta = 1) = p + \delta$ for a fixed $\delta \in (-p, 1-p)$. It is called counterfactual because it may assign $D_\delta = 1$ to an individual whose $D = 0$, and viceversa. As δ varies over $[-p, 1-p]$, we obtain a collection of counterfactual policies which is denoted by $\mathscr{D}$. Somewhat casually, we also call the collection $\mathscr{D}$ a *sequence of policies*. When a particular counterfactual policy D_δ belongs to $\mathscr{D}$ we write $D_\delta \in \mathscr{D}$. The counterfactual outcome we would observe for a given $D_\delta \in \mathscr{D}$ is

$$Y_{D_\delta} = r(D_\delta, X, U),$$

where we implicitly assumes that the potential outcomes are not affected by the manipulation of D.[2]

Remark 1.1. *Suppose D indicates union status, and $Y(0)$ and $Y(1)$ represent the potential (log) wages for non-unionized and unionized individuals. A sequence of policies $\mathscr{D}$ can be thought of as a manipulation of the union status of individuals: we can expand the union to cover $p+\delta$ individuals, for $\delta > 0$. The resulting counterfactual outcome Y_{D_δ} can be the union wage for nonunionized worker. However, the distribution of the potential wages may well depend on the proportion of unionized individuals. For example, we expect that the average wage for unionization sector when everyone is unionized to be different from the case when half of the workers are unionized. Card et al. (2004) make this point explicit by indexing the moments of the distribution of wages by proportion of unionized workers. Thus, we may write $Y(0;p)$ and $Y(1;p)$, where $p = \Pr(D=1)$. This is a violation of the Stable Unit Value Treatment Assumption (SUTVA, Imbens and Rubin (2015)), which is not allowed in our setting.*

We will evaluate the effect of a counterfactual policy on the quantiles of the distribution of the outcome. In particular, we look at two quantities: the global and the marginal effects. Let $F_Y^{-1}(\tau)$ and $F_{Y_{D_\delta}}^{-1}(\tau)$ denote the τ-quantiles of Y and Y_{D_δ} respectively.[3]

Definition 1.1 (Global and Marginal Effects)**.** *For a given sequence of policies $\mathscr{D}$, the uncondi-*

[2] Strictly speaking, the counterfactual outcome Y_{D_δ} is not well defined until we define $\mathscr{D}$, the collection of counterfactual policies.

[3] Alternatively, we use y_τ instead of $F_Y^{-1}(\tau)$, and $y_{\tau,\delta}$ instead of $F_{Y_{D_\delta}}^{-1}(\tau)$.

tional global effect at the τ-quantile is

$$G_{\tau,D_\delta} := F^{-1}_{Y_{D_\delta}}(\tau) - F^{-1}_Y(\tau),$$

and the unconditional marginal effect at the τ-quantile is

$$M_{\tau,\mathscr{D}} := \lim_{\delta \to 0} \frac{F^{-1}_{Y_{D_\delta}}(\tau) - F^{-1}_Y(\tau)}{\delta}$$

whenever this limit exists.

The global effect G_{τ,D_δ} is the comparison of quantiles of the counterfactual distribution vs. the observed distribution. For example, it could tell us what could happen to the median under a particular policy D_δ. The marginal effect $M_{\tau,\mathscr{D}}$ can be interpreted as an ordinary derivative: for small δ, it provides an approximation to the direction of the change in a given τ-quantile. The bigger the absolute value of the marginal effect, the stronger is the effect of a small change in the proportion of treated individuals.

Figure 1.1 contains a graphical representation of the marginal and global effects for a given quantile. The solid blue line is the map $\delta \mapsto F^{-1}_{Y_{D_\delta}}(\tau)$ for a given τ and a given sequence of policies.[4] The slope of the red line, which is the tangent of $\delta \mapsto F^{-1}_{Y_{D_\delta}}(\tau)$ at $\delta = 0$ is the unconditional marginal effect M_τ. For a given δ_0, the green bar parallel to the Y axis is the unconditional global effect $G_{\tau,D_{\delta_0}} = F^{-1}_{Y_{D_{\delta_0}}}(\tau) - F^{-1}_Y(\tau)$. Implicit is the fact that for $\delta = 0$, $F^{-1}_{Y_{D_0}}(\tau) = F^{-1}_Y(\tau)$.

The marginal effect on the unconditional quantiles of an outcome was first studied by Firpo et al. (2009). The identification arguments of Firpo et al. (2009) are based on a distributional invariance assumption: the distribution of the outcome for the original treatment group (under the original policy regime) is the same as that for the new treatment group (under the new policy regime), and this also holds for the control groups under the two policy regimes.[5] Rothe (2012)

[4]This map is generally unknown, but its derivative at $\delta = 0$ is easier to identify.

[5]See the proof to Corollary 3 of the working paper version Firpo et al. (2007).

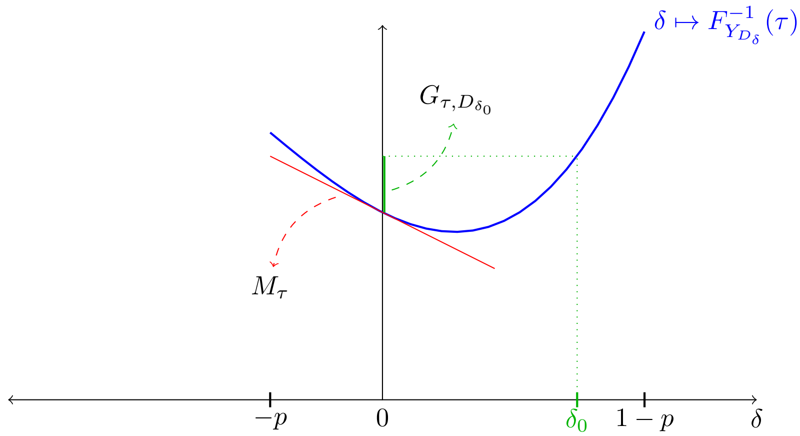

Figure 1.1. Global and marginal effects.

provides a general treatment for functionals of the unconditional distribution of the outcome. What we call a global effect, Rothe (2012) refers to as a *Fixed Partial Policy Effect*, and what we call a marginal effect, Rothe (2012) refers to as a *Marginal Partial Distributional Policy*.

Before we proceed, we will settle the question of existence of the marginal effect. Theorem 1.1 provides a set of general sufficient conditions.

Theorem 1.1 (Existence of Marginal Effect)**.** *Consider a sequence of policies $\mathscr{D}$ such that*

1. $F_{Y_{D_0}}(y) = F_Y(y)$ *for any* $y \in \mathscr{Y}$*;*

2. *The map* $\delta \mapsto F_{Y_{D_\delta}}(y)$ *is differentiable at* $\delta = 0$ *uniformly in* $y \in \mathscr{Y}$*, with derivative* $\dot{F}_{Y,\mathscr{D}}(y)$*, that is*

$$\lim_{\delta \downarrow 0} \sup_{y \in \mathscr{Y}} \left| \frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} - \dot{F}_{Y,\mathscr{D}}(y) \right| = 0;$$

3. *The map* $y \mapsto \dot{F}_{Y,\mathscr{D}}(y)$ *is continuous at* $F_Y^{-1}(\tau)$*.*

Then, $M_{\tau,\mathscr{D}}$ *exists and is given by*

$$M_{\tau,\mathscr{D}} = -\frac{\dot{F}_{Y,\mathscr{D}}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))}.$$

The conditions and the proof of this Theorem come from viewing the marginal effect as a Hadamard derivative. The first condition, $F_{Y_{D_0}}(y) = F_Y(y)$, is particular to this setting, though. The requirement $F_{Y_{D_0}}(y) = F_Y(y)$ states that for $D_0 \in \mathscr{D}$, the limiting counterfactual distribution F_{D_0} matches the observed distribution F_Y. This might not necessarily be the case. Indeed, we could define a marginal effect with respect to $F_{Y_{D_0}}$ instead which would avoid the "discontinuity" at $\delta = 0$. However, this would be of limited interest. The next example illustrates a case where $F_{Y_{D_0}}(y) \neq F_Y(y)$.

Example 1.1 (Threshold Crossing Model)**.** *Suppose that individuals select into treatment by $D = \mathbb{1}\{V \leq 0.5\}$ for $V \sim U_{[0,1]}$. Consider the sequence of policies $D_\delta = \mathbb{1}\{V \leq 0.5+\delta\}$, and $\delta \geq 0$. Then,*

$$F_{Y_{D_\delta}}(y) = \Pr(Y \leq y | V \leq 0.5+\delta)(0.5+\delta) + \Pr(Y \leq y | V > 0.5+\delta)(1-0.5-\delta).$$

In this case, $F_{Y_{D_0}} = F_Y$. However, if the sequence of policies is $D_\delta = \mathbb{1}\{V > 0.5-\delta\}$, then $F_{Y_{D_0}}$ might not coincide with F_Y.

The second condition in Theorem 1.1, that of uniform differentiability of the map $\delta \mapsto F_{Y_{D_\delta}}(y)$, is more abstract. Essentially, it requires that small departures from 0 to $\delta > 0$ should not induce large (uniform) changes in the counterfactual distribution $F_{Y_{D_\delta}}$.

1.2.1 Double Unconfoundedness

Usually, in order to identify different parameters of interest, an unconfoundedness assumptions is imposed: $D \perp Y(0), Y(1) \| X$. This means that, given a sufficiently rich set of covariates, assignment to treatment is *as good as random*. That is, assignment is independent of potential outcomes. Kaplan (2020) notes, however, that such requirement falls short when dealing with counterfactual policies and, additionally, $D_\delta \perp Y(0), Y(1) \| X$ is necessary as well. Using the structural model in (1.1), we can equivalently write this assumption as $D, D_\delta \perp U \| X$.

Assumption 1.2 (Double Unconfoundedness)**.** *For the model given in* (1.1)*, the following holds*

1. $D \perp U \| X$;

2. *for every* $D_\delta \in \mathscr{D}$, $D_\delta \perp U \| X$;

3. $\mathscr{X}_d$ *is the common support of* $X|D = d$ *and* $X|D_\delta = d$;

Remark 1.2. *The first part of Assumption 1.2, namely,* $D \perp U \| X$ *is fundamentally untestable. However, the second part,* $D_\delta \perp U \| X$ *depends on the policy maker. One way to achieve this, if covariates are discrete, is for the policy maker to randomized* D_δ *by strata. Finally, the common support assumption is made mostly for notational convenience. A sufficient condition for it to be satisfied is that the support of* $X|D = 0$ *equals the support of* $X|D = 1$*. This is typically an assumption that is made in order to avoid extrapolation: for any given value of* X*, that is, for comparable units, we can find some that were assigned to treatment, and some that were assigned to control.*

The usefulness of Assumption 1.2 is that is allows us to point identify the counterfactual distribution $F_{Y_{D_\delta}}$. This counterfactual distribution is the basis for the global and marginal effects that we are interested in.

Lemma 1.1. *Under Assumption 1.2, the counterfactual distribution* $F_{Y_{D_\delta}}(y)$ *is*

$$\begin{aligned} F_{Y_{D_\delta}}(y) = (p+\delta) \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) dF_{X|D_\delta=1}(x) \\ + (1-p-\delta) \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) dF_{X|D_\delta=0}(x). \end{aligned}$$

The counterfactual distribution, which is a marginal distribution, is given by the usual reweighting of conditional distributions. However, we integrate against the *new* conditional distribution of the covariates $X|D_\delta$.

The result in Lemma 1.1 allows us to immediately identify the global effect. We just need to invert the counterfactual distribution, and compare this to the inverse of the observed distribution of the outcome. The situation with respect to the marginal effect deserves special attention due to the requirements for Theorem 1.1 to hold. We need the uniform convergence of the derivative over the support $\mathscr{Y}$, which is the support of both Y and Y_{D_δ}. To see what this entails, we can write the result in Lemma 1.1 as a quotient:

$$\begin{aligned}
\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} &= p \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) d\left(\frac{F_{X|D_\delta=1}(x) - F_{X|D=1}(x)}{\delta}\right) \\
&+ (1-p) \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) d\left(\frac{F_{X|D_\delta=0}(x) - F_{X|D=0}(x)}{\delta}\right) \\
&+ \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) dF_{X|D_\delta=1}(x) \\
&- \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) dF_{X|D_\delta=0}(x)
\end{aligned}$$

According to Theorem 1.1, we need to take the uniform limit in $y \in \mathscr{Y}$ as $\delta \to 0$. An important quantity involved in this limit is the propensity score, denoted by $P(x) := \Pr(D = 1|X = x)$. For a given sequence of counterfactual policies, we have a corresponding sequence of counterfactual propensity scores $P_\delta(x) := \Pr(D_\delta = 1|X = x)$. The map $\delta \mapsto P_\delta(x)$ will play an important role later. The counterfactual propensity scores have to satisfy the consistency restriction

$$E[\Pr(D_\delta = 1|X)] = E[\Pr(D = 1|X)] + \delta = p + \delta. \tag{1.2}$$

The following assumptions are needed to obtain the marginal effect.

Assumption 1.3.

1. *For every $D_\delta \in \mathscr{D}$, $\mathscr{Y}$ is the common support of Y and Y_{D_δ};*

2. *For every $D_\delta \in \mathscr{D}$, and $d = 0, 1$, $F_{X|D_\delta=d}(x)$ and $F_{X|D=d}(x)$ are absolutely continuous with respect to Lebesgue measure, with densities given by $f_{X|D_\delta=d}(x)$ and $f_{X|D=d}(x)$ respectively;*

3. *For $d = 0,1$, and every $x \in \mathscr{X}_d$,*

$$\lim_{\delta \to 0} f_{X|D_\delta = d}(x) = f_{X|D=d}(x);$$

4. *For every $D_\delta \in \mathscr{D}$ and every $x \in \mathscr{X}_d$,*

$$\frac{P_\delta(x) - P(x)}{\delta} \leq m(x)$$

with $E[|m(X)|] < \infty$, and

$$\dot{P}(x) := \left. \frac{\partial P_\delta(x)}{\partial \delta} \right|_{\delta=0}$$

exists.

5. *The map $(y,x) \mapsto F_{Y|D=d,X=x}(y)$ is continuous for $d = 0,1$;*

Remark 1.3. *Most of these assumptions are used to justify taking the limit under the integral sign. The existence of $\dot{P}(x)$ is very important for our results and we will analyze its interpretation more in detail below. In later sections, the existence of $\dot{P}(x)$ will follow from the form of the selection equation (a threshold crossing equation) and the type of counterfactual policy considered. When covariates are discrete, the domination assumption by the function $m(x)$ can be dispensed of, since we just have a finite sum, under which it is permissible to pass the limit.*

Theorem 1.2. *Under Assumptions 1.1, 1.2 and 1.3, the marginal effect at the τ-quantile exists, and is given by*

$$M_{\tau,\mathscr{D}} = -\frac{E\left[\left(F_{Y|D=1,X}(F_Y^{-1}(\tau)) - F_{Y|D=0,X}(F_Y^{-1}(\tau))\right)\dot{P}(X)\right]}{f_Y(F_Y^{-1}(\tau))}.$$

When $\dot{P}(x) \equiv 1$, this is precisely the estimand Firpo et al. (2009) consider. However, their derivation, as seen in Corollary 3 of Firpo et al. (2007), does not include covariates. Instead, Firpo et al. (2007) assumes what we call *distributional invariance*: for $d = 0,1$, $F_{Y_{D_\delta}|D_\delta=d}(y) = F_{Y|D=d}(y)$. Using this, we obtain

$$
\begin{aligned}
F_{Y_{D_\delta}}(y) &= (p+\delta)F_{Y_{D_\delta}|D_\delta=1}(y) + (1-p-\delta)F_{Y_{D_\delta}|D_\delta=0}(y) \\
&= (p+\delta)F_{Y|D=1}(y) + (1-p-\delta)F_{Y|D=0}(y) \\
&= F_Y(y) + \delta\left(F_{Y|D=1}(y) - F_{Y|D=0}(y)\right),
\end{aligned}
$$

where the last line uses the decomposition $F_Y(y) = pF_{Y|D=1}(y) + (1-p)F_{Y|D=0}(y)$. Now,

$$
\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} = F_{Y|D=1}(y) - F_{Y|D=0}(y)
$$

which implies trivially that $\dot{F}_{Y,\mathscr{D}}(y) = F_{Y|D=1}(y) - F_{Y|D=0}(y)$. Note further that $\dot{F}_{Y,\mathscr{D}}$ is independent of $\mathscr{D}$. Thus, the result of Theorem 1.2 is a generalization of Firpo et al. (2009), not only for the inclusion of the covariates, but also for the inclusion of the adjustment term $\dot{P}(x)$.

We refer to $\dot{P}(x)$ as the *derivative of the propensity score* and we will analyze it more carefully in later sections. We can write

$$
\dot{P}(x) = \lim_{\delta \to 0} \frac{\Pr(D_\delta = 1|X=x) - \Pr(D=1|X=x)}{\delta}
$$

The only term "in control" of the policy maker is $\Pr(D_\delta = 1|X=x)$. Suppose the policy maker has some weighting function $\omega(x)$ to allocate the newly treated, such that

$$
\Pr(D_\delta = 1|X=x) - \Pr(D=1|X=x) = \delta\omega(x)
$$

and $E[\omega(X)] = 1$. In this case, in turns out that $\dot{P}(x) = \omega(x)$. This sheds light on the interpretation of the derivative of the propensity score. When $\dot{P}(x) \equiv 1$, it means that all "subpopulations" are weighted equally by the policy maker when considering a marginal expansion of the treated proportion.

Example 1.2. *A counterfactual policy that weights every subpopulation equally is one where* $\Pr(D_\delta = 1|X = x) = \Pr(D = 1|X = x) + \delta$. *This counterfactual policy trivially satisfies the consistency requirement in* (1.2)*. Here,* $\dot{P}(x) \equiv 1$.

Example 1.3. *Suppose that X is discrete, taking values in* $\mathscr{X} = \{1, 2, \ldots, k\}$*. Consider the following counterfactual policy*

$$\Pr(D_\delta = 1|X = x) = \Pr(D = 1|X = x) + \frac{\delta}{k}\frac{1}{\Pr(X = x)}.$$

which gives less weight to the bigger subpopulations, as measured by $\Pr(X = x)$*. Followiing* (1.2)*, we have*

$$\sum_{x=1}^{k} \Pr(D_\delta = 1|X = x)\Pr(X = x) = p + \sum_{x=1}^{k} \frac{\delta}{k}\frac{1}{\Pr(X = x)}\Pr(X = x) = p + \delta$$

and $\dot{P}(x) = (k\Pr(X = x))^{-1}$*. The marginal effect is then*

$$M_{\tau,\mathscr{D}} = -\frac{1}{kf_Y(F_Y^{-1}(\tau))}\sum_{x=1}^{k}\left[F_{Y|D=1,X=x}(F_Y^{-1}(\tau)) - F_{Y|D=0,X=x}(F_Y^{-1}(\tau))\right].$$

Example 1.4. *Again, suppose that* $\mathscr{X} = \{1, 2, \ldots, k\}$*, and that* $\Pr(X = 1) \leq \Pr(X = 2) \leq \ldots \leq \Pr(X = k)$*. The policy*

$$\Pr(D_\delta = 1|X = x) = \Pr(D = 1|X = x) + \delta\frac{x}{E[X]}$$

allocates more weight to subpopulations with bigger values of x. In this case, $\dot{P}(x) = x/E[X]$*, and the marginal effect is*

$$M_{\tau,\mathscr{D}} = -\frac{\sum_{x=1}^{k}\left[F_{Y|D=1,X=x}(F_Y^{-1}(\tau)) - F_{Y|D=0,X=x}(F_Y^{-1}(\tau))\right]x\Pr(X = x)}{E[X]f_Y(F_Y^{-1}(\tau))}.$$

These examples highlight the fact that the marginal effects can be quite different depending on the type of policies considered. The derivative of the propensity score $\dot{P}(x)$ is the term that reflects this adjustment. Intuitively, $\dot{P}(x)$ reflects the direction in which we depart from the status quo policy in the counterfactual policy. When the adjustment given by the derivative of the propensity score is ignored, that is, when we incorrectly set $\dot{P}(x) \equiv 1$, the estimand will exhibit asymptotic bias given by

$$\text{Bias}_\tau = \frac{E\left[\left(F_{Y|D=1,X}(F_Y^{-1}(\tau)) - F_{Y|D=0,X}(F_Y^{-1}(\tau))\right)(\dot{P}(X)-1)\right]}{f_Y(F_Y^{-1}(\tau))}.$$

In the next section we are going to analyze the bias by explicitly allowing both ignoring $\dot{P}(x)$ and departing from the double unconfoundedness assumption.

1.2.2 Threshold Crossing Model

In some cases, it is possible that even after controlling for all the observed covariates, some correlation remains between D and U. This is the case, for example, when D follows a threshold crossing equation, and dislike or resistance to treatment is correlated with the potential outcomes. In this section we add some more structure to the selection equation in a manner that the double unconfoundedness assumption does not hold. In doing so, we will obtain an expression for the asymptotic bias in this case.

Following Heckman and Vytlacil (1999, 2001a, 2005), we assume that selection into treatment is determined by the threshold-crossing equation

$$D = \mathbb{1}\{V \le \mu(X)\}, \tag{1.3}$$

where $\mu(X)$ can be regarded as the benefit from the treatment and V as the cost of the treatment. Individuals decide to take up the treatment if and only if its benefit outweighs its cost. Alternatively, we can think of $\mu(X)$ as the utility and V as the disutility from participating in the program. While we observe (D, X, Y), we observe neither $U := (U_0, U_1)$ nor V. Also, we do not restrict the dependence among U,X, and V, hence, D could be endogenous.

Note that when we condition on X in the selection equation (1.3), D depends on V which is allowed to be correlated with U. Thus, in this case, unconfoundedness will not hold.

Recall the propensity score is $P(x) := \Pr(D=1|X=x)$. In the threshold crossing model, it has a convenient representation. Indeed

$$P(x) = \Pr(V \leq \mu(X)|X=x) = F_{V|X}(\mu(x)|x).$$

If the conditional CDF $F_{V|X}(\cdot|x)$ is a strictly increasing function for all $x \in \mathscr{X}$, the support of X, we have

$$D = \mathbb{1}\{V \leq \mu(X)\} = \mathbb{1}\left\{F_{V|X}(V|X) \leq F_{V|X}(\mu(X)|X)\right\} = \mathbb{1}\{U_D \leq P(X)\},$$

where $U_D = F_{V|X}(V|X)$. Furthermore, U_D is uniform on $[0,1]$ and is independent of X.[6]

As before, we let

$$r(D,X,U) = (1-D)\,r_0(X,U_0) + Dr_1(X,U_1),$$

then we have $Y = r(D,X,U)$.

In the previous section, we consider policies that switches individuals from one group the other or viceversa. Here, we consider a policy intervention that changes the propensity score

[6]To see this, we denote $F_{V|X}(v|x) := \Pr(V \leq v|X=x)$ by $G_x(v)$. We use the notation $G_x(v)$ when we view $F_{V|X}(V|X)$ as a function of v for a given x. We have

$$\begin{aligned}
\Pr(U_D \leq u|X=x) &= \Pr\left(F_{V|X}(v|x) \leq u|X=x\right) = \Pr\left(F_{V|X}(v|x) \leq u|X=x\right) \\
&= \Pr(G_x(V) \leq u|X=x) = \Pr\left(V \leq G_x^{-1}(u)|X=x\right) \\
&= G_x\left[G_x^{-1}(u)\right] = u.
\end{aligned}$$

By the law of iterated expectations, we have $\Pr(U_D \leq u) = u$. Also, $\Pr(U_D \leq u|X=x)$ does not depend on x, so U_D is indeed uniform on $[0,1]$ and is independent of X. Note that, in general, if U_D is a deterministic function of both V and X, then U_D is not independent of (V,X), but this does not rule out the possibility that U_D is independent of X. Conditioning on X, U_D is a deterministic function of V only, and hence U_D and V are dependent conditioning on X. See also Heckman and Vytlacil (1999).

from $P(x)$ to $P_\delta(x)$. Our manipulation of the propensity will be done via a manipulation of the selection equation. We can do so because we assume a particular form of the selection equation. The drawback is that now we do not know who are the individuals that will take treatment. Naturally, the counterfactual propensity score will satisfy the consistency requirement as in (1.2). That is

$$E\left[P_\delta(X)\right] = E\left[P(X)\right] + \delta.$$

Note that the consistency requirement does not always imply that $P_0(x) = P(x)$. However, the particular form of the policies analyzed in this section will ensure this. Under the new policy, our model becomes

$$\begin{aligned} Y_\delta &= r\left(D_\delta, X, U\right) = \left(1 - D_\delta\right) Y\left(0\right) + D_\delta Y\left(1\right), \\ D_\delta &= \mathbb{1}\left\{U_D \leq P_\delta(X)\right\}. \end{aligned} \tag{1.4}$$

For notational convenience, when $\delta = 0$, we drop the subscript, and we write Y, D, and $P(X)$ for Y_0, D_0, and $P_0(X)$, respectively. It is important to highlight that, regardless of the value of δ, the dependence pattern between U and V given X is the same. More precisely, the conditional distribution of (U,V) given X is invariant to the value of δ. Equivalently, the conditional distribution of (U,U_D) given X under $\delta = \delta_o$ for any δ_o is the same as that under $\delta = 0$.

Example 1.5. *Many different policies can be used to change the propensity score. As an example, we can increase everyone's benefit function by the same amount $s_\mu(\delta)$ with the normalization that $s_\mu(0) = 0$. In this case, we have*

$$P_\delta(x) = \Pr\left[V \leq \mu\left(X\right) + s_\mu\left(\delta\right) | X = x\right] = F_{V|X}(\mu(x) + s_\mu\left(\delta\right) | x), \tag{1.5}$$

and $P_0(x) = P(x)$. Effectively, we induce a location change in the benefit function (or the cost *function) while keeping intact the dependence between U and V (or U_D) given X.*

In this section we are concerned with a generic policy that changes the propensity score from $P(x)$ to $P_\delta(x)$. We are agnostic about how the policy intervention achieves this change; we impose only the consistency restriction that $E[P_\delta(X)] = p+\delta$, and that $P_0(x) = P(x)$ for every $x \in \mathscr{X}$.

In order to find the unconditional quantile treatment effect, we first make a support assumption.

Assumption 1.4. ***Support Assumption*** *For $d = 0,1$, the support of $Y(d)$ given (U_D, X) does not depend on (U_D, X).*

Note that D_δ is a function of (U_D, X). Under the above assumption, the support of $Y(d)$ given D_δ also does not depend on D_δ. We denote the support of $Y(d)$ by $\mathscr{Y}(d)$, which is the common support regardless of whether we condition on (U_D, X). It is also the common support regardless of whether we condition on D_δ.

For some $\varepsilon > 0$, define $N_\varepsilon := \{\delta : |\delta| \le \varepsilon\}$. To simplify notation, let y_τ and $y_{\tau,\delta}$ denote the τ-quantiles of Y and Y_{D_δ} respectively.[7] For every $\delta \in N_\varepsilon$, we have

$$\begin{aligned} F_{Y_\delta}(y_{\tau,\delta}) &= (p+\delta)\Pr\left(Y(1) \le y_{\tau,\delta} | D_\delta = 1\right) \\ &+ (1-p-\delta)\Pr\left(Y(0) \le y_{\tau,\delta} | D_\delta = 0\right) \\ &= \int_{\mathscr{Y}(1)} \mathbb{1}\left\{y \le y_{\tau,\delta}\right\}(p+\delta) f_{Y(1)|D_\delta}(y|1)dy \\ &+ \int_{\mathscr{Y}(0)} \mathbb{1}\left\{y \le y_{\tau,\delta}\right\}(1-p-\delta) f_{Y(0)|D_\delta}(y|0)dy, \end{aligned} \tag{1.6}$$

where we have used the support assumption so that the support of $Y(d)$ given $D_\delta = d$ is still $\mathscr{Y}(d)$.

The goal is to linearize $(p+\delta) f_{Y(1)|D_\delta}(y|1)$ and $(1-p-\delta) f_{Y(0)|D_\delta}(y|0)$ around $\delta = 0$. To this end, we make some technical assumptions.

[7] We also denote these as $F_Y^{-1}(\tau)$ and $F_{Y_{D_\delta}}^{-1}(\tau)$ respectively.

Assumption 1.5. ***Regularity Conditions***

(a) *For $d = 0, 1$, the random variables $(Y(d), U_D, X)$ are absolutely continuous with joint density given by $f_{Y(d),U_D|X} f_X$.*

(b) *(i) For $d = 0, 1$, $u \mapsto f_{Y(d)|U_D,X}(y|u,x)$ is continuous for almost all $y \in \mathscr{Y}(d)$ and almost all $x \in \mathscr{X}$. (ii) For $d = 0, 1$, for almost all $y \in \mathscr{Y}(d)$,*

$$\sup_{x\in\mathscr{X}} \sup_{\delta\in N_\varepsilon} f_{Y(d)|U_D,X}(y|P_\delta(x), x) < \infty.$$

(c) *(i) $p = E[P(X)] \in (0,1)$. (ii) For each $x \in \mathscr{X}$, the map $\delta \mapsto P_\delta(x)$ is continuously differentiable on N_ε. (iii) $\sup_{x\in\mathscr{X}} \sup_{\delta\in N_\varepsilon} \left| \frac{\partial P_\delta(x)}{\partial \delta} \right| < \infty$.*

Assumption 1.6. ***Domination Conditions*** *For $d = 0, 1$,*

$$\begin{aligned} \int_{\mathscr{Y}(d)} \sup_{\delta\in N_\varepsilon} f_{Y(d)|D_\delta}(y|d) dy &< \infty, \\ \int_{\mathscr{Y}(d)} \sup_{\delta\in N_\varepsilon} \left| \frac{\partial f_{Y(d)|D_\delta}(y|d)}{\partial \delta} \right| dy &< \infty. \end{aligned}$$

In Assumption 1.5 the supremum over $x \in \mathscr{X}$ can be replaced by the essential supremum over $x \in \mathscr{X}$.

Lemma 1.2. *Let Assumptions 1.4 and 1.5 hold. For $d = 0, 1$, the map $\delta \mapsto f_{Y(d)|D_\delta}(y|d)$ is continuously differentiable on N_ε for almost all $y \in \mathscr{Y}(d)$.*

Using Lemma 1.2, we expand $F_{Y_\delta}(y)$ in (1.6) around $\delta = 0$ to obtain the approximation in the following lemma.

Lemma 1.3. *Let Assumptions 1.4–1.6 hold. Then*

$$\begin{aligned} F_{Y_\delta}(y) &= F_Y(y) + \delta E\left[\left\{F_{Y(1)|U_D,X}(y|P(X), X) - F_{Y(0)|U_D,X}(y|P(X), X)\right\} \dot{P}(X)\right] \\ &+ o(|\delta|), \end{aligned}$$

uniformly over $y \in \mathscr{Y} := \mathscr{Y}(0) \cup \mathscr{Y}(1)$ *as* $\delta \to 0$*, where*

$$\dot{P}(X) := \left. \frac{\partial P_{\delta}(X)}{\partial \delta} \right|_{\delta=0}.$$

Remark 1.4. *Lemma 1.3 provides a linear approximation to* F_{Y_δ}*, the CDF of the outcome variable under* D_δ*. Essentially, it says that the proportion of individuals with outcome below y under the new selection rule, that is,* $F_{Y_\delta}(y)$*, will be equal to the proportion of individuals with outcome below y under the original selection rule, that is,* $F_Y(y)$*, plus an adjustment given by the marginal entrants. Consider* $\delta > 0$ *and* $P_\delta(x) > P(x)$ *for all* $x \in \mathscr{X}$ *as an example. In this case, because of the policy intervention, the individuals who are on the margin, namely those with* $u_D = P(x)$*, will switch their treatment status from* 0 *to* 1*. Such a switch contributes to* $F_{Y_\delta}(y)$ *by the amount* $F_{Y(1)|U_D,X}(y|P(x),x) - F_{Y(0)|U_D,X}(y|P(x),x)$*, averaged over the distribution of X for a certain subpopulation. We will show later that the subpopulation is exactly the group of individuals who are on the margin under the existing policy regime.*

Remark 1.5. *Existence of the derivative of the propensity score* $\dot{P}(x)$ *is guaranteed by Assumption 1.5.(c).*

Remark 1.6. *The linear approximation to* F_{Y_δ} *is uniform in* $\mathscr{Y}$ *as* $\delta \to 0$*. We need the uniform approximation because below we "invert" the approximation to obtain the quantiles.*

Theorem 1.3. *Let Assumptions 1.4–1.6 hold. Assume further that* $f_Y(y_\tau) > 0$*. Then*

$$\begin{aligned} M_{\tau,\mathscr{D}} &:= \lim_{\delta \to 0} \frac{y_{\tau,\delta} - y_\tau}{\delta} \\ &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(0) \le y_\tau\right\} | U_D = P(x), X = x\right] \dot{P}(x) f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(1) \le y_\tau\right\} | U_D = P(x), X = x\right] \dot{P}(x) f_X(x)\, dx. \end{aligned} \tag{1.7}$$

Theorem 1.3 shows that, among the individuals with $X = x$, only those for whom $u_D = P(x)$ will contribute to the marginal unconditional quantile effect. Among the group defined by $X=x$, there is a subgroup of individuals who are indifferent between participating and not participating: those for whom $u_D = P(x)$, that is, those for whom v satisfies $F_{V|X}(V|x) = P(x)$. A small incentive will induce a change in the treatment status only for this subgroup of individuals. It is the change in the treatment status, and hence the change in the composition of $Y(1)$ and $Y(0)$ in the observed outcome Y, that changes the unconditional quantiles of Y.

Theorem 1.3 shows that the unconditional quantile effect depends also on $\dot{P}(X)$. Under Assumption 1.5(c), we have

$$\int_{\mathscr{X}} \frac{\partial P_\delta(x)}{\partial \delta} f_X(x) = \frac{\partial}{\partial \delta} \int_{\mathscr{X}} P_\delta(x) f_X(x) = \frac{\partial}{\partial \delta}(p+\delta) = 1$$

for any $\delta \in N_\varepsilon$, hence $E[\dot{P}(X)] = 1$. Thus the integrals in (1.7) can be regarded as a weighted mean with the weight given by $\dot{P}(x)$. Note that $\dot{P}(x)$ depends on how we choose to modify the propensity score, that is, it depends on who the marginal entrants are. Different propensity score interventions can result in different sets of marginal entrants and different unconditional quantile effects. As in the case of double unconfoundedness, intuitively, the derivative of the propensity $\dot{P}(x)$ indicates the "direction" of departure from the status quo policy.

For intuition on this, consider the case where $\delta > 0$ and $P_\delta(x) \geq P(x)$ for all $x \in \mathscr{X}$. Then we have

$$\begin{aligned} \dot{P}(x) &= \lim_{\delta \to 0} \frac{\Pr(U_D \leq P_\delta(x)) - \Pr(U_D \leq P(x))}{\delta} \\ &= \lim_{\delta \to 0} \frac{\Pr(P(X) < U_D \leq P_\delta(X) | X = x)}{\delta}. \end{aligned}$$

Thus, $\dot{P}(x)$ measures the relative contribution to the overall improvement in the participation rate (i.e., δ) for the individuals with $X=x$. For each value of x, only individuals on the margin ("the marginal individuals") will change their treatment status and contribute to the overall improvement in the participation rate. The relative "thickness" of the margin depends on x and is measured by $\dot{P}(x)$.

We can use Figure 1.2 to convey the intuition behind $\dot{P}(X)$. The figure illustrates the marginal individuals under the existing and new policy regimes. The marginal individuals are those with $u_D = P_\delta(x)$. Under the existing policy regime, the marginal individuals lie on the 45-degree line in the $(P(x), u_D)$-plane. For easy reference, we call it the marginal curve, which is the set of points $\{(P(x), u_D) : u_D = P(x)\}$. Under the new policy regime, the marginal curve is now $\{(P(x), u_D) : u_D = P_\delta(x)\}$. Note that we can rewrite $u_D = P_\delta(x)$ as $u_D = P(x) + [P_\delta(x) - P(x)]$. Thus the new marginal curve can be obtained by shifting every point on the original marginal curve up by $P_\delta(x) - P(x)$. The magnitude of the upward shift is approximately $\dot{P}(x)\,\delta$, which is, in general, different for different values of x. The integral of the difference of the two marginal curves (i.e., the area of the gray region) weighted by the marginal density $f_X(\cdot)$ of X is equal to δ.[8]

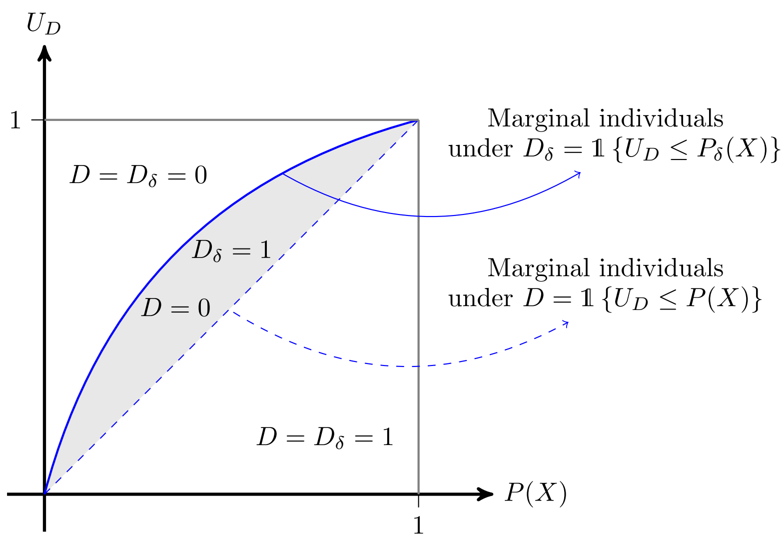

Figure 1.2. Marginal individuals under different policies.

To understand the weight $f_X(x)\dot{P}(x)$ that appears in Theorem 1.3, let ε be a small positive number. Then, $f_X(x)\varepsilon$ measures the proportion of individuals for whom X is in

[8]Note that the two marginal curves coincide in the limit as $\delta \to 0$, and so in the limit we can define the marginal individuals as those with $U_D = P(x)$. In our discussions, "the marginal individuals" may refer to the group of individuals with $U_D = P(x)$ or the group of individuals with $U_D = P_\delta(x)$. Which group we refer to should be clear from the context.

$[x-\varepsilon/2, x+\varepsilon/2]$. Note that for $X \in [x-\varepsilon/2, x+\varepsilon/2]$, the propensity scores under D and D_δ are approximately $P(x)$ and $P_\delta(x)$. The proportion of the individuals for whom $X \in [x-\varepsilon/2, x+\varepsilon/2]$ and who have switched their treatment status from 0 to 1 is then equal to

$$f_X(x) \cdot [P_\delta(x) - P(x)] \cdot \varepsilon.$$

Scaling this by δ, which is the overall proportion of the individuals who have switched the treatment status, we obtain $(f_X(x) \cdot [P_\delta(x) - P(x)]/\delta) \cdot \varepsilon$. Thus $f_X(x) \cdot [P_\delta(x) - P(x)]/\delta$ can be regarded as the density function of X among those who have switched the treatment status from 0 to 1 as a result of the policy intervention.

Mathematically, we have

$$\begin{aligned} &\frac{\Pr(X \in [x-\varepsilon/2, x+\varepsilon/2] \,|D=0, D_\delta = 1)}{\varepsilon} \\ &= \frac{\Pr(X \in [x-\varepsilon/2, x+\varepsilon/2], D=0, D_\delta=1)}{\varepsilon \cdot \Pr(D=0, D_\delta=1)} \\ &= \frac{\Pr(X \in [x-\varepsilon/2, x+\varepsilon/2], U_D \in [P(x), P_\delta(x)])}{\varepsilon \cdot \delta}, \end{aligned}$$

so taking the limit as $\varepsilon \to 0$, we obtain

$$\lim_{\varepsilon \to 0} \frac{\Pr(X \in [x-\varepsilon/2, x+\varepsilon/2] \,|D=0, D_\delta=1)}{\varepsilon} = f_X(x) \frac{P_\delta(x) - P(x)}{\delta}.$$

Thus $f_X(x)[P_\delta(x) - P(x)]/\delta$ is the density of X among those who respond positively to the policy intervention, that is, those with $D=0$ and $D_\delta = 1$. Graphically, $f_X(x)[P_\delta(x) - P(x)]/\delta$ is the conditional density of X conditional on $(P(x), U_D)$ being in the gray region in Figure 1.2.

Letting $\delta \to 0$, we obtain

$$\lim_{\delta \to 0} f_X(x) \frac{P_\delta(x) - P(x)}{\delta} = f_X(x) \dot{P}(x).$$

That is, $f_X(x)\dot{P}(x)$ is the limit of the density of X among those with $D=0, D_\delta=1$. We can therefore refer to $f_X(x)\dot{P}(x)$ as the density of the distribution of X over the marginal subpopulation that consists of all marginal individuals.

In view of the above interpretation of $f_X(x)\dot{P}(x)$, Theorem 1.3 shows that the unconditional quantile effect is equal to the change in the influence functions for the marginal individuals, weighted by the density of the distribution of X over those marginal individuals.

Noting that $f_X(x)$ is the density of the distribution of X over the entire population, we can regard $\dot{P}(x)$ as the Radon–Nikodym (RN) derivative of the subpopulation distribution with respect to the population distribution. Even if $\dot{P}(x)$ is not positive for all $x \in \mathscr{X}$, the Radon–Nikodym interpretation is still valid. In this case, the distribution with density $f_X(x)\dot{P}(x)$ with respect to the Lebesgue measure is a signed measure.

The next corollary decomposes the unconditional quantile effect into an apparent component, which neglects the adjustment given by $\dot{P}(X)$ and a bias component.

Corollary 1.4. *Let the assumptions in Theorem 1.3 hold. Then*

$$M_{\tau,\mathscr{D}} = A_\tau - B_\tau,$$

where[9]

$$\begin{aligned} A_\tau &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\{Y \le y_\tau\} \mid D=0, X=x\right] f_X(x)\,dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\{Y \le y_\tau\} \mid D=1, X=x\right] f_X(x)\,dx, \end{aligned} \tag{1.8}$$

and[10]

$$B_\tau = B_{1\tau} + B_{2\tau},$$

[9] The apparent term does not depend on the sequence $\mathscr{D}$.
[10] The bias term depends on the sequence $\mathscr{D}$.

for

$$\begin{aligned} B_{1\tau} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}}\left[F_{Y(0)|D,X}(y_\tau|0,x)-F_{Y(1)|D,W}(y_\tau|1,x)\right]\left[1-\dot{P}(x)\right]f_X(x)\,dx \\ &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}}\left[F_{Y|D,X}(y_\tau|1,x)-F_{Y|D,X}(y_\tau|0,x)\right]\dot{P}(x)\,f_X(x)\,dx, \\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}}\left[F_{Y|D,X}(y_\tau|1,x)-F_{Y|D,X}(y_\tau|0,x)\right]f_X(x)\,dx \end{aligned}$$

and

$$\begin{aligned} B_{2\tau} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}}\left[F_{Y(0)|D,X}(y_\tau|0,x)-F_{Y(0)|U_D,X}(y_\tau|P(x),x)\right]\dot{P}(x)\,f_X(x)\,dx \\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}}\left[F_{Y(1)|D,X}(y_\tau|1,x)-F_{Y(1)|U_D,X}(y_\tau|P(x),x)\right]\dot{P}(x)\,f_X(x)\,dx. \end{aligned} \tag{1.9}$$

To facilitate understanding of Corollary 1.4, we can define and organize the average influence functions (AIF) in a table:

Table 1.1. Average Influence Functions

	Difference of Average Influence Functions
U_D	$E_x[\psi_\tau(Y(1))-\psi_\tau(Y(0))\,|U_D=P(X)]$
D	$E_x[\psi_\tau(Y(1))\,|D=1]-E_x[\psi_\tau(Y(0))\,|D=0]$

Notes: The differences in each case are AIF for $Y(1)$ minus AIF for $Y(0)$.

where

$$\psi_\tau(Y(d))=\frac{\tau-\mathbb{1}\{Y(d)\le y_\tau\}}{f_Y(y_\tau)}.$$

In the above, $E_x[\cdot]$ stands for the conditional mean operator given $X=x$. For example

$$E_x[\psi_\tau(Y(0))\,|D=0]=E[\psi_\tau(Y(0))\,|D=0,X=x].$$

Let

$$\begin{aligned}\psi_{\Delta,U_D}(x) &:= E_x[\psi_\tau(Y(1)) - \psi_\tau(Y(0))|U_D = P(x)],\\ \psi_{\Delta,D}(x) &:= E_x[\psi_\tau(Y(1))|D=1] - E_x[\psi_\tau(Y(0))|D=0].\end{aligned}$$

The unconditional quantile effect $M_{\tau,\mathscr{D}}$ is the average of the difference $\psi_{\Delta,U_D}(x)$ with respect to the distribution of X over the marginal subpopulation. The average *apparent* effect A_τ is the average of the difference $\psi_{\Delta,D}(x)$ with respect to the distribution of X over the whole population distribution. It is also equal to the limit of the unconditional quantile estimator of Firpo et al. (2009), where the endogeneity of the treatment selection is ignored. The discrepancy between $M_{\tau,\mathscr{D}}$ and A_τ gives rise to the asymptotic *bias* B_τ :

$$\begin{aligned}B_\tau &= A_\tau - M_{\tau,\mathscr{D}} = E\left[\psi_{\Delta,D}(X)\right] - E\left[\psi_{\Delta,U_D}(X)\dot{P}(X)\right]\\ &= \underbrace{E\left\{\psi_{\Delta,D}(X)\left[1-\dot{P}(X)\right]\right\}}_{B_{1\tau}} + \underbrace{E\left\{\left[\psi_{\Delta,D}(X) - \psi_{\Delta,U_D}(X)\right]\dot{P}(X)\right\}}_{B_{2\tau}}. \qquad (1.10)\end{aligned}$$

It is easy to show that the $B_{2\tau}$ given here is identical to that given in (1.9).

Equation (1.10) decomposes the asymptotic bias into two components. The first one, $B_{1\tau}$, captures the heterogeneity of the averaged apparent effects averaged over two different subpopulations. For every x, $\psi_{\Delta,D}(X)$ is the average effect of D on $[\tau - \mathbb{1}\{Y \leq y_\tau\}]/f_Y(y_\tau)$ for the individuals with $X = x$. These effects are averaged over two different distributions of X: the distribution of X for the marginal subpopulation (i.e., $\dot{P}(x) f_X(x)$) and the distribution of X for the whole population (i.e., $f_X(x)$). $B_{1\tau}$ is equal to the difference of these two average effects. If the effect $\psi_{\Delta,D}(x)$ does not depend on x, then $B_{1\tau} = 0$. If $\dot{P}(x) = 1$, then the distribution of X over the whole population is the same as that over the subpopulation, and hence $B_{1\tau} = 0$ in that case as well.

The second bias component, $B_{2\tau}$, has a difference-in-differences interpretation. Each of $\psi_{\Delta,D}(\cdot)$ and $\psi_{\Delta,U_D}(\cdot)$ is the difference in the average influence functions associated with the counterfactual outcomes $Y(1)$ and $Y(0)$. However, $\psi_{\Delta,D}(\cdot)$ is the difference over the two subpopulations who actually choose $D=1$ and $D=0$, while $\psi_{\Delta,U_D}(\cdot)$ is the difference over the marginal subpopulation. So $\psi_{\Delta,D}(\cdot)-\psi_{\Delta,U_D}(\cdot)$ is a difference in differences. $B_{2\tau}$ is simply the average of this difference in differences with respect to the distribution of X over the marginal subpopulation. This term arises because the change in the distributions of Y for those with $D=1$ and those with $D=0$ is different from that for those whose U_D is just above $P(x)$ and those whose U_D is just below $P(x)$. Thus we can label $B_{2\tau}$ as a *marginal selection bias*.

If $\psi_{\Delta,D}(x)=\psi_{\Delta,U_D}(x)$ for almost all $x\in\mathscr{X}$, then $B_{2\tau}=0$. The condition $\psi_{\Delta,D}(x)=\psi_{\Delta,U_D}(x)$ is the same as

$$\begin{aligned}&E_x\left[\psi_\tau(Y(1))-\psi_\tau(Y(0))\,|U_D=P(x)\right]\\&=E_x\left[\psi_\tau(Y(1))\,|D=1\right]-E_x\left[\psi_\tau(Y(0))\,|D=0\right].\end{aligned}$$

Equivalently,

$$\begin{aligned}&E_x\left[\psi_\tau(Y(1))\,|U_D=P(x)\right]-E_x\left[\psi_\tau(Y(1))\,|D=1\right]\\&\qquad=E_x\left[\psi_\tau(Y(0))\,|U_D=P(w)\right]-E_x\left[\psi_\tau(Y(0))\,|D=0\right].\end{aligned}$$

The condition resembles the parallel-paths assumption or the constant-bias assumption in a difference-in-differences analysis. If U_D is independent of (U_0,U_1) given X, then this condition holds and $B_{2\tau}=0$.

In general, when U_D is not independent of (U_0,U_1) given X, and X enters the selection equation, we have $B_{1\tau}\neq 0$ and $B_{2\tau}\neq 0$, hence $M_{\tau,\mathscr{D}}\neq A_\tau$. If $\dot{P}(X)$ is not identified, then $B_{1\tau}$ is not identified. In general, $B_{2\tau}$ is not identified without an instrument. Therefore, without an instrument, the asymptotic bias can not be eliminated and Π_τ is not identified.

It is not surprising that in the presence of endogeneity, the unconditional quantile estimator of Firpo et al. (2009) is asymptotically biased. The virtue of Corollary 1.4 is that it provides a closed-form characterization and clear interpretations of the asymptotic bias. To the best of our knowledge, this bias formula is new in the literature. From a broad perspective, the asymptotic bias B_τ is the unconditional quantile counterpart of the endogenous bias of the OLS estimator in a linear regression framework.

The bias decomposition is not unique. Corollary 1.4 gives only one possibility. We can also write

$$B_\tau = \underbrace{E\left\{\psi_{\Delta,U_D}(X)\left[1-\dot{P}(X)\right]\right\}}_{\tilde{B}_{1\tau}} + \underbrace{E\left[\psi_{\Delta,D}(X) - \psi_{\Delta,U_D}(X)\right]}_{\tilde{B}_{2\tau}}.$$

The interpretations of $\tilde{B}_{1\tau}$ and $\tilde{B}_{2\tau}$ are similar to those of $B_{1\tau}$ and $B_{2\tau}$ with obvious and minor modifications. The non-uniqueness of the decomposition when two or more quantities change simultaneously is well expected.[11]

Example 1.6. *Consider a setting of full independence: $V \perp U_0 \perp U_1 \perp X$ (i.e., every subset is independent of its complement). In this case, $B_{2\tau} = 0$ and by equation (1.7), the UQE is*

$$\begin{aligned} M_{\tau,\mathscr{D}} &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(0) \le y_\tau\right\}|X=x\right] \dot{P}(X) f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(1) \le y_\tau\right\}|X=x\right] \dot{P}(X) f_X(x)\, dx. \end{aligned}$$

Following (1.8), the apparent effect is

$$\begin{aligned} A_\tau &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(0) \le y_\tau\right\}|X=x\right] f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(1) \le y_\tau\right\}|X=x\right] f_X(x)\, dx. \end{aligned}$$

In general, we will still have a bias term given by $B_{1\tau}$ unless $\dot{P}(x) = 1$ or the difference $E\left[\mathbb{1}\left\{Y(0) \le y_\tau\right\}|X=x\right] - E\left[\mathbb{1}\left\{Y(1) \le y_\tau\right\}|X=x\right]$ does not depend on x. In general, both conditions fail if the covariate X enters both the outcome equations and the selection equation. In this case, the usual unconditional quantile regression estimator will be asymptotically biased.

[11] Here, $\psi_{\Delta,U_D}(\cdot)$ changes to $\psi_{\Delta,D}(\cdot)$ and $\dot{P}(x) f_X(x)$ changes to $f_X(x)$.

On the other hand, under full independence, there will be no asymptotic bias if the treatment has a constant effect across the values of the covariates (i.e., $E[\mathbb{1}\{Y(0)\le y_\tau\}|X=x]-E[\mathbb{1}\{Y(1)\le y_\tau\}|X=x]$ *does not depend on x) or if the distribution of X over the whole population is the same as that over the marginal subpopulation. Note that if X does not enter the outcome equation, then* $E[\mathbb{1}\{Y(0)\le y_\tau\}|X=x]$ *is equal to* $E[\mathbb{1}\{Y(1)\le y_\tau\}|X=x]$ *under the condition* $(U_0,U_1)\perp X$*. As a result, there will be no bias. If X does not enter the selection equation so that* $\mu(x)=\mu_0$ *for a constant* μ_0*, then* $\dot{P}(x)=1$*, and the distribution of X over the whole population is the same as that over the marginal subpopulation. As a result, there will be no bias in that case, either.*[12]

Remark 1.7. *We can relate the result of Example 1.6 to the results under double unconfoundedness. In Example 1.6, the marginal effect (there is no bias) is*

$$\begin{aligned} A_\tau &= \frac{1}{f_Y(y_\tau)}\int_{\mathcal{X}} E[\mathbb{1}\{Y(0)\le y_\tau\}|X=x]\, f_X(x)\,dx \\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathcal{X}} E[\mathbb{1}\{Y(1)\le y_\tau\}|X=x]\, f_X(x)\,dx. \end{aligned}$$

By Assumption 1.2, $D\perp U\|X$*, which implies*

$$E[\mathbb{1}\{Y(0)\le y_\tau\}|X=x]=E[\mathbb{1}\{Y\le y_\tau\}|X=x,D=0],$$

and

$$E[\mathbb{1}\{Y(1)\le y_\tau\}|X=x]=E[\mathbb{1}\{Y\le y_\tau\}|X=x,D=1].$$

Using this, we get

$$A_\tau=-\frac{E\left[\left(F_{Y|D=1,X}(y_\tau)-F_{Y|D=0,X}(y_\tau)\right)\right]}{f_Y(y_\tau)},$$

[12]To see this, note that when $P_\delta(x)$ does not depend on x, we have that $P_\delta=p+\delta$, which implies that $\dot{P}=1$.

which is the result of Theorem 1.2 for the case $\dot{P}(x) \equiv 1$.

1.2.3 UQE under Location Shift in the Cost or Benefit Function

In this subsection, we modify the propensity score by inducing a location shift in the cost or benefit function. Note that an increase in the benefit has the same effect as reducing the cost by the same amount. It is innocuous to focus on a location shift in the benefit function, which is one of many ways to change the propensity score. From the perspective of policy design, we ask and address the following question: given the dependence between $U = (U_0, U_1)$ and V in the population, how will the unconditional quantiles of Y change if we manage to improve the benefit from participating in the program by $s_\mu(\delta)$ for each individual in the population? While U and V are dependent, the incremental change $s_\mu(\delta)$ is the *same* for all individuals and hence is exogenous.

Recall that $P_\delta(x)$ was given in (1.5):

$$P_\delta^\mu(x) = \Pr\left[V \leq \mu(x) + s_\mu(\delta) | X = x\right] = F_{V|X}(\mu(x) + s_\mu(\delta) | x). \tag{1.11}$$

To emphasize that the change is induced on $\mu(\cdot)$, we have added a subscript "μ" to $P_\delta(x)$. We will use $P_\delta^\mu(x)$ exclusively for this case. Note that $s_\mu(\delta)$ is determined implicitly by the equation $E\left[P_\delta^\mu(X)\right] = p + \delta$. The following lemma characterizes how $s_\mu(\delta)$ and $P_\delta^\mu(x)$ will change in response to a small change in δ.

Lemma 1.4. *Assume that (i) (V, X) are absolutely continuous random variables with joint density $f_{V,X}(v,x)$ given by $f_{V|X}(v|x) f_X(x)$; (ii) $f_{V|X}(v|x)$ is continuous in v for almost all $x \in \mathscr{X}$; (iii) $\int_{\mathscr{X}} \sup_{\delta \in N_\varepsilon} f_{V|X}(\mu(x) + s_\mu(\delta) | x) f_X(x) dx < \infty$; and (iv) $\int_{\mathscr{X}} f_{V|X}(\mu(x) + s_\mu(\delta) | x) f_X(x) dx \neq 0$*

for all $\delta \in N_\varepsilon$. Then for all $\delta \in N_\varepsilon$,

$$\begin{aligned}\frac{\partial s_\mu(\delta)}{\partial \delta} &= \frac{1}{\int_{\mathscr{X}} f_{V|X}(\mu(x)+s_\mu(\delta)|x)f_X(x)dx},\\ \frac{\partial P^\mu_\delta(x)}{\partial \delta} &= \frac{f_{V|X}(\mu(x)+s_\mu(\delta)|x)}{\int_{\mathscr{X}} f_{V|X}(\mu(x)+s_\mu(\delta)|x)f_X(x)dx}.\end{aligned}$$

In particular,

$$\begin{aligned}\left.\frac{\partial s_\mu(\delta)}{\partial \delta}\right|_{\delta=0} &= \frac{1}{\int_{\mathscr{X}} f_{V|X}(\mu(x)|x)f_X(x)dx},\\ \dot{P}^\mu(x) = \left.\frac{\partial P^\mu_\delta(x)}{\partial \delta}\right|_{\delta=0} &= \frac{f_{V|X}(\mu(x)|x)}{\int_{\mathscr{X}} f_{V|X}(\mu(x)|x)f_X(x)dx}.\end{aligned}$$

Combining Theorem 1.3 and Lemma 1.4, we obtain a representation of the unconditional quantile effect under a location shift in the benefit function. We use $\Pi_{\tau,\mu}$ to denote this effect.

Corollary 1.5. *Let the assumptions in Theorem 1.3 hold. Then under the location shift in the benefit function given by (1.11), we have*

$$\begin{aligned}\Pi_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(0)\le y_\tau\}|U_D=P(x),X=x\right]\dot{P}^\mu(x)f_X(x)dx\\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(1)\le y_\tau\}|U_D=P(x),X=x\right]\dot{P}^\mu(x)f_X(x)dx, \qquad (1.12)\end{aligned}$$

where

$$\dot{P}^\mu(x) = \left.\frac{\partial P^\mu_\delta(x)}{\partial \delta}\right|_{\delta=0} = \frac{f_{V|X}(\mu(x)|x)}{\int_{\mathscr{X}} f_{V|X}(\mu(x)|x)f_X(x)dx}.$$

To bring our setting closer to that of Firpo et al. (2009) and shed more light on Theorem 1.3, we consider the special case where $\mu(x)=\mu_0$ for a constant μ_0 and V is independent of X. In this case, the selection is based on unobservables only. We have

$$P^\mu_\delta(x) = F_{V|X}(\mu_0+s_\mu(\delta)|x) = F_V(\mu_0+s_\mu(\delta)),$$

which does not depend on x. In particular, $P_0^{\mu}(X) = F_V(\mu_0) = p$. The selection rule becomes $D = \mathbb{1}\{U_D \leq p\}$, where U_D is uniform on $[0,1]$. Using Lemma 1.4, we find that

$$\dot{P}^{\mu} = \left.\frac{\partial P_{\delta}^{\mu}(x)}{\partial \delta}\right|_{\delta=0} = \frac{f_V(\mu_0)}{\int_{\mathscr{X}} f_{V|X}(\mu_0|x) f_X(x)dx} = \frac{f_V(\mu_0)}{\int_{\mathscr{X}} f_V(\mu_0) f_X(x)dx} = 1.$$

In fact, we can also obtain this directly from $P_{\delta}^{\mu}(x) = p + \delta$. It then follows that the distribution of X over the whole population is the same as that over the marginal subpopulation. As a result, one of the asymptotic bias terms, namely $B_{1\tau}$, disappears.

Using Theorem 1.3 and Corollary 1.4, we obtain the following corollary.

Corollary 1.6. *Let the assumptions in Theorem 1.3 hold. If $\mu(x) = \mu_0$ for some constant μ_0 and V is independent of X, then*

$$\Pi_{\tau,\mu} = \frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\{Y(0) \leq y_\tau\} - \mathbb{1}\{Y(1) \leq y_\tau\} | U_D = p\right]. \tag{1.13}$$

In addition, $\Pi_{\tau,\mu} = A_{\tau,\mu} - B_{\tau,\mu}$ *for*

$$\begin{aligned} A_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)} \int E\left[\mathbb{1}\{Y \leq y_\tau\} | D=0, X=x\right] f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int E\left[\mathbb{1}\{Y \leq y_\tau\} | D=1, X=x\right] f_X(x)\, dx \qquad (1.14) \\ &= \frac{1}{f_Y(y_\tau)} \left(E\left[\mathbb{1}\{Y \leq y_\tau\} | D=0\right] - E\left[\mathbb{1}\{Y \leq y_\tau\} | D=1\right]\right) \qquad (1.15) \end{aligned}$$

and

$$\begin{aligned} B_{\tau,\mu} = &\frac{1}{f_Y(y_\tau)} \left[F_{Y(0)|D}(y_\tau|0) - F_{Y(0)|U_D}(y_\tau|p)\right] \\ &- \frac{1}{f_Y(y_\tau)} \left[F_{Y(1)|D}(y_\tau|1) - F_{Y(1)|U_D}(y_\tau|p)\right]. \end{aligned}$$

The representation of $\Pi_{\tau,\mu}$ in (1.13) shows that the unconditional quantile effect is equal to the difference of the influence functions averaged over the conditional distribution of the outcome variable given $U_D = p$. That is, the unconditional quantile effect is driven by the individuals for whom $U_D = p$. These individuals are *ex-ante* indifferent between choosing $D = 1$ and $D = 0$. They are on the margin—a small increase in the benefit will push them to switch from non-participating (i.e., $D = 0$) to participating (i.e., $D = 1$). It is the switch that changes the outcome of interest, and hence its unconditional distribution and unconditional quantiles.

If $U_D \perp U$, then $F_{Y(d)|D}(y|d) = F_{Y(d)|U_D}(y|p) = F_{Y(d)}(y)$ for any $y \in \mathscr{Y}(d)$ and for $d = 0$ and $d = 1$. As a result, $B_\tau = 0$. Thus the asymptotic bias B_τ vanishes in the absence of endogeneity. This is, of course, well expected.

1.2.4 Examples of Asymptotic Bias

In the presence of endogeneity, it is not easy to evaluate or even sign the asymptotic bias. In general, the joint distribution of (U, U_D) given the covariate X is needed for this purpose. This is not atypical. For a nonlinear estimator such as the unconditional quantile estimator, its asymptotic properties often depend on the full data generating process in a non-trivial way. This is in sharp contrast with a linear estimator such as the OLS in a linear regression model whose properties depend on only the first few moments of the data. Next, we present some examples in which we can derive the asymptotic bias explicitly.

Example 1.7. ***Non-uniformity of Endogeneity Bias across Quantiles***

Consider the model

$$
\begin{aligned}
Y(d) &= U \in \mathbb{R}, \text{ for } d = 0, 1, \\
Y &= DY(1) + (1 - D)Y(0), \\
D &= \mathbb{1}\{V \leq 0\},
\end{aligned}
$$

where U and V are correlated. Noting that $Y(0) = Y(1)$, the treatment has no effect on the outcome, and so

$$\mathbb{1}\{Y(0) \leq y_\tau\} - \mathbb{1}\{Y(1) \leq y_\tau\} = 0.$$

As a result, $\Pi_{\tau,\mu} = 0$*. By Corollary 1.6,* $A_{\tau,\mu} = B_{\tau,\mu}$*, that is, the estimator of Firpo et al. (2009) is an estimator of the asymptotic bias only. In this case, since* $Y = U$*, we have*

$$B_{\tau,\mu} = \frac{1}{f_U(u_\tau)} \left[F_{U|D}(u_\tau|0) - F_{U|D}(u_\tau|1) \right],$$

where u_τ *is the* τ*-quantile of* U*:* $F_U(u_\tau) = \tau$*. This shows that the sign of the asymptotic bias depends on* $F_{U|D}(u_\tau|0) - F_{U|D}(u_\tau|1)$*. In the presence of endogeneity, the two distribution functions* $F_{U|D}(\cdot|0)$ *and* $F_{U|D}(\cdot|1)$ *are not the same. Unless one distribution function first-order stochastically dominates the other, it is necessarily true that* $B_{\tau,\mu}$ *is positive for some quantile levels and negative for others.*

Figure 1.3 shows a case where the bias is positive for higher quantiles and negative for lower quantiles. Thus it is not sufficient to use the sign of the correlation between U *and* V *to sign the asymptotic bias for the unconditional quantile effect at all quantile levels.*

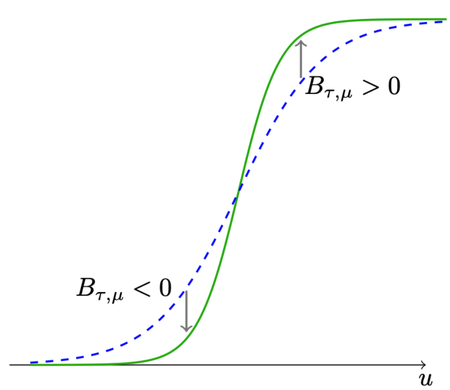

Figure 1.3. Non-uniform bias.

Example 1.8. ***Asymptotic Bias with Exogenous Treatment***

Consider the model

$$\begin{aligned} Y(0) &= q(X) + U_0, \\ Y(1) &= q(X) + \beta + U_1, \\ D &= \mathbb{1}\{V \le \mu(x)\}, \\ Y &= (1-D)Y(0) + DY(1), \end{aligned}$$

where $q(\cdot)$ *and* $\mu(\cdot)$ *are functions of* X*, and* X *is independent of* (U_0, U_1, V)*. By Theorem 1.3, and assuming that* $f_{U_0|V} = f_{U_1|V}$*, we have*

$$\Pi_{\tau,\mu} = \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[\int_{y_\tau - q(x) - \beta}^{y_\tau - q(x)} f_{U_0|V}(u|\mu(X))du \right] \tilde{f}_X(x)dx$$

where

$$\tilde{f}_X(x) = \frac{f_V(\mu(x)) f_X(x)}{\int_{\mathscr{X}} f_V(\mu(\check{x})) f_X(\check{x}) d\check{x}}.$$

It follows from Corollary 1.4 that the apparent effect is

$$\begin{aligned} A_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \frac{\int_{\mu(X)}^{\infty} \left[\int_{-\infty}^{y_\tau - q(x)} f_{U_0|V}(u|v)du \right] f_V(v)\, dv}{1 - F_V(\mu(x))} f_X(x)dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \frac{\int_{-\infty}^{\mu(x)} \left[\int_{-\infty}^{y_\tau - q(x) - \beta} f_{U_1|V}(u|v)du \right] f_V(v)\, dv}{F_V(\mu(x))} f_X(x)dx. \end{aligned}$$

Details of this and the expression for $f_Y(y_\tau)$ *can be found in the supplementary appendix.*

To compute $\Pi_{\tau,\mu}$ *and* $A_{\tau,\mu}$ *numerically, we set* $\beta = 2$*,* $q(x) = \mu(x) = e^x$*, and we assume that* X *is standard normal and that* (U_0, U_1, V) *is normal with mean* 0 *and variance*

$$\Sigma = \begin{bmatrix} 1 & 0 & \rho \\ 0 & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix},$$

where ρ is the correlation between U_0 and V, and between U_1 and V. Different values of ρ lead to different degrees of endogeneity. When $\rho = 0$, the treatment selection is exogenous.

Figure 1.4 plots the asymptotic bias $B_{\tau,\mu} := A_{\tau,\mu} - \Pi_{\tau,\mu}$ as a function of τ for $\rho = 0, .25, .5, .75, .9$. As in Example 1.7, we can see that the asymptotic bias is not uniform across quantiles. Hence, any attempt to sign the bias based on the "sign" or degree of the endogeneity (i.e., the sign or magnitude of ρ) is futile.

It is intriguing to see that, even in the case of exogenous treatment selection (i.e., $\rho = 0$), we still have an asymptotic bias. When $\rho = 0$, the asymptotic bias from the second source is 0, as

$$\begin{aligned} F_{Y(0)|D,X}(y_\tau|0,x) &= F_{Y(0)|U_D,X}(y_\tau|P(X),x) = F_{Y(0)|X}(y_\tau|x), \\ F_{Y(1)|D,X}(y_\tau|1,x) &= F_{Y(1)|U_D,X}(y_\tau|P(X),x) = F_{Y(1)|X}(y_\tau|x). \end{aligned}$$

The asymptotic bias from the first source is

$$\begin{aligned} B_{1\tau,\mu} &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[F_{Y|D,X}(y_\tau|1,x) - F_{Y|D,X}(y_\tau|0,x)\right] \dot{P}(X) f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[F_{Y|D,X}(y_\tau|1,x) - F_{Y|D,X}(y_\tau|0,x)\right] f_X(x)\, dx. \end{aligned}$$

In this example, $\dot{P}(x) \neq 1$ and

$$F_{Y|D,X}(y_\tau|0,x) - F_{Y|D,X}(y_\tau|1,x) = \Pr(e^x + U_0 < y_\tau) - \Pr(e^x + \beta + U_1 < y_\tau) \neq 0$$

for all $\beta \neq 0$. So the asymptotic bias from the first source, $B_{1\tau,\mu}$, is not equal to 0 when $\beta \neq 0$.

Why is there a bias when the treatment is exogenous? To improve the overall participation rate from p to $p + \delta$, we change the benefit function by the same amount $s_\mu(\delta)$, that is, from $\mu(x)$ to $\mu(x) + s_\mu(\delta)$. Depending on the value of x, such a change in the benefit function will have a differential effect on the treatment rate. In other words, $P_\delta(x) - P(x)$ and hence $\dot{P}(x)$, *will depend on x. As a result, the distribution of X over the whole population will not be the same as that over the marginal subpopulation. This difference creates a wedge between the unconditional quantile effect* $\Pi_{\tau,\mu}$ *and the average apparent effect* $A_{\tau,\mu}$ *when the "apparent" difference* $F_{Y|D,X}(y_\tau|0,x) - F_{Y|D,X}(y_\tau|1,x)$ *depends on x.*

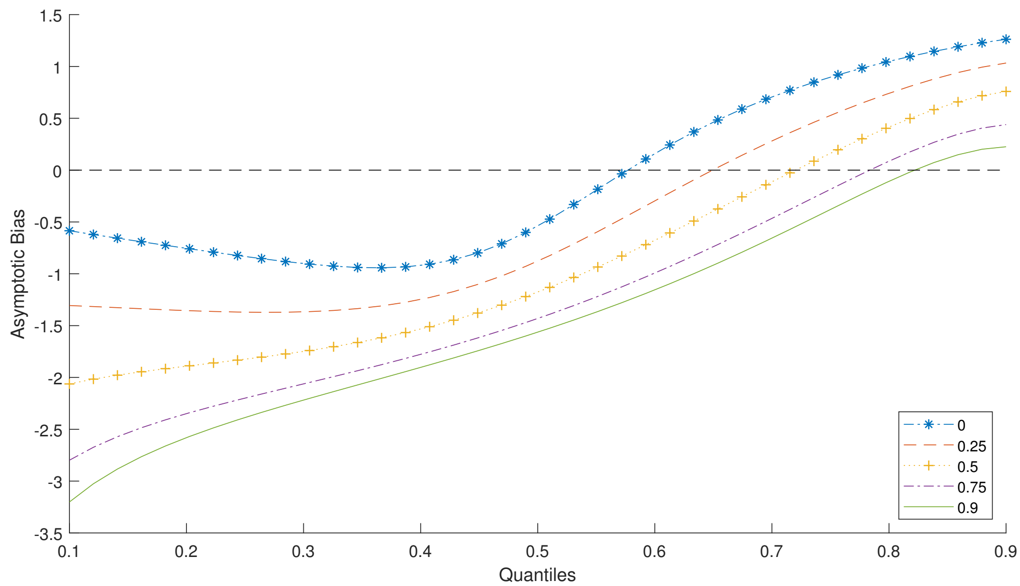

Figure 1.4. Asymptotic bias for $\rho = 0, 0.25, 0.5, 0.75, 0.9$.

1.3 Continuous Target Variable Case

In this case, the model is $Y = r(X, \tilde{W}, U)$, and X is a target variable with absolutely continuous distribution function F_X. In this case, the thought experiment is that the policy maker successfully manipulates X to achieve X^*, and obtain a counterfactual outcome $Y^* = r(X^*, \tilde{W}, U)$. The goal is to characterize what happens to the unconditional distribution of the outcome. The distribution Y is a transformation of the *joint* distribution of $(X, \tilde{W}, U)$. Thus, the unconditional distribution of Y^* is not well-defined until we do not specify the *new* joint distribution of $(X^*, \tilde{W}, U)$.

Heuristically, an identification argument runs as follows:

$$F_{Y^*}(y) = \int_{\mathscr{X}^*} F_{Y^*|X^*=x}(y) dF_{X^*}(x)$$

by the law of iterated expectations. Suppose that $F_{Y^*|X^*=x}(y) = F_{Y|X=x}(y)$, and that the supports $\mathscr{X}^* = \mathscr{X}$, then

$$F_{Y^*}(y) = \int_{\mathscr{X}} F_{Y|X=x}(y) dF_{X^*}(x)$$

and we have identified $F_{Y^*}(y)$. This is, basically, the identification strategy of Firpo et al. (2009).

In this section, we try to disentangle exactly what is behind the distributional assumption $F_{Y^*|X^*=x}(y) = F_{Y|X=x}(y)$, and what are the biases associated when it is not satisfied. We will also be more careful when dealing with the support $\mathscr{X}^*$ and how it relates to $\mathscr{X}$.

Assumption 1.7 (Support)**.** *The support of the target variable X is* $[x_l, x_u] \subset [-\infty, \infty]$*, and its density vanishes at the boundary:* $f_X(x_l) = f_X(x_u) = 0$*.*

We will consider counterfactual policies in the form of a location shift. To this end, recall that, for some $\varepsilon > 0$, we defined $N_\varepsilon := \{\delta : |\delta| \leq \varepsilon\}$.

Assumption 1.8 (Location Shift)**.** *For any* $\varepsilon > 0$*, and* $\delta \in N_\varepsilon$*,* $X^* = X + \delta$*.*

We will focus on the marginal effect, which is defined as

$$M_\tau = \lim_{\delta \to 0} \frac{F_{Y^*}^{-1}(\tau) - F_Y^{-1}(\tau)}{\tau}$$

whenever this limit exists, where $Y^* = r(X + \delta, \tilde{W}, U)$ under Assumption 1.8.

We would like to characterize the counterfactual distribution $F_{Y^*}(y)$. Since we are interested in X^*, we will define $\tilde{U} := (\tilde{W}, U)'$, and denote its support by $\tilde{\mathscr{U}}$. We have

$$F_{Y^*}(y) = \int_{x_l+\delta}^{x_u+\delta} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x, \tilde{u}) \leq y\} f_{X^*, \tilde{U}}(x, \tilde{u}) d\tilde{u} dx.$$

Assumption 1.9 (Conditional Density)**.** *The joint density of* $(X,\tilde{U})$ *is given by* $f_{\tilde{U},X}(u,x) = f_{\tilde{U}|X=x}(u) f_X(x)$.

Under assumption 1.9 we can find the joint density of $(X^*,\tilde{U})$. We have

$$\Pr(X^* \le x, \tilde{W} \le w, U \le u) = \Pr(X \le x - \delta, \tilde{W} \le w, U \le u),$$

so that for $x \in (x_l + \delta, x_u + \delta)$, we have

$$f_{X^*,\tilde{U}}(x,\tilde{u}) = f_{\tilde{U},X}(\tilde{u}, x-\delta) = f_{\tilde{U}|X=x-\delta}(\tilde{u}) f_X(x-\delta).$$

Therefore, we can write

$$\begin{aligned} F_{Y^*}(y) &= \int_{x_l+\delta}^{x_u+\delta} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f_X(x-\delta) d\tilde{u} dx \\ &= \int_{x_l}^{x_u} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x+\delta,\tilde{u}) \le y\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u} dx \end{aligned}$$

Using the indicator function to keep track of the limits of integration we have

$$F_{Y^*}(y) = \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} \mathbb{1}\{x_l \le x - \delta \le x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f_X(x-\delta) d\tilde{u} dx$$

and

$$F_Y(y) = \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} \mathbb{1}\{x_l \le x \le x_u\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u} dx.$$

So, the gap between $F_{Y^*}(y)$ and $F_Y(y)$ can be decomposed as

$$
\begin{aligned}
& F_{Y^*}(y) - F_Y(y) \\
&= \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f_X(x-\delta) d\tilde{u}dx \\
&- \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x \leq x_u\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx \\
&= \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} \\
&\times \left[f_{\tilde{U}|X=x-\delta}(\tilde{u}) f_X(x-\delta) - f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) \right] d\tilde{u}dx \\
&+ \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \left[\mathbb{1}\{x_l \leq x-\delta \leq x_u\} - \mathbb{1}\{x_l \leq x \leq x_u\} \right] f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx.
\end{aligned}
$$

The second term, the one involving the difference $\left[\mathbb{1}\{x_l \leq x-\delta \leq x_u\} - \mathbb{1}\{x_l \leq x \leq x_u\} \right]$ is exactly zero when $x_l = -\infty$ and $x_u = \infty$. The first term can be further decomposed to obtain three terms that explain the gap.

$$
\begin{aligned}
& F_{Y^*}(y) - F_Y(y) \\
&= \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) \left[f_X(x-\delta) - f_X(x) \right] d\tilde{u}dx \\
&+ \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_X(x) \left[f_{\tilde{U}|X=x-\delta}(\tilde{u}) - f_{\tilde{U}|X=x}(\tilde{u}) \right] d\tilde{u}dx \\
&+ \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \left[\mathbb{1}\{x_l \leq x-\delta \leq x_u\} - \mathbb{1}\{x_l \leq x \leq x_u\} \right] \\
&\times f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx.
\end{aligned}
$$

The next step is to do a Taylor expansion around $\delta = 0$, and provided rigorous conditions under which the remainder, which is integrated, is of order $o(|\delta|)$. We write

$$
F_{Y^*}(y) - F_Y(y) = \mathscr{I}_1(y) + \mathscr{I}_2(y) + \mathscr{I}_3(y) \tag{1.16}
$$

and we treat each term at a time. We need two additional assumptions. One concerning smoothness of densities, and one concerning domination.

Assumption 1.10 (Differentiable densities)**.** *1. The density $f_X(x)$ is continuously differentiable in (x_l, x_u), with derivative given by $f'_X(x)$;*

2. The map $x \mapsto f_{\tilde{U}|X=x}(\tilde{u})$ is continuously differentiable in (x_l, x_u) for every $\tilde{u} \in \mathscr{W} \times \mathscr{U}$, with derivative with respect to x given by $f'_{\tilde{U}|X=x}(\tilde{u})$.

Assumption 1.11 (Domination)**.** *The following domination conditions hold*

$$\int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \sup_{\delta \in N_\varepsilon} \left[f_{\tilde{U}|X=x-\delta}(\tilde{u}) \right] \sup_{\delta' \in N_\varepsilon} |f'_X(x-\delta')| d\tilde{u} dx < \infty, \tag{1.17}$$

and

$$\int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \sup_{\delta \in N_\varepsilon} \left[f'_{\tilde{U}|X=x-\delta}(\tilde{u}) \right] f_X(x) d\tilde{u} dx < \infty. \tag{1.18}$$

Lemma 1.5. *Under Assumptions 1.7, 1.8, 1.9, 1.10, and 1.11*

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \frac{F_{Y^*}(y) - F_Y(y)}{\delta} - \dot{F}(y) \right| = 0$$

where

$$\dot{F}_Y(y) = E\left[\left. \frac{\partial F_{Y|X=x}(y)}{\partial x} \right|_{x=X} \right] - E\left[\mathbb{1}\left\{ r(X, \tilde{U}) \leq y \right\} \frac{\partial}{\partial x} \log \left. \frac{f_{X,\tilde{U}}(x, \tilde{u})}{f_X(x) f_{\tilde{U}}(\tilde{u})} \right|_{x=X, \tilde{u}=\tilde{U}} \right]. \tag{1.19}$$

Corollary 1.7. *The marginal effect at the τ-quantile is then*

$$\begin{aligned} M_\tau = & \underbrace{-\frac{1}{f_Y(y_\tau)} E\left[\left. \frac{\partial F_{Y|X=x}(y_\tau)}{\partial x} \right|_{x=X} \right]}_{:=A_\tau\text{: Estimand of Firpo et al. (2009)}} \\ & + \underbrace{\frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\left\{ r(X, \tilde{U}) \leq y_\tau \right\} \frac{\partial}{\partial x} \log \left. \frac{f_{X,\tilde{U}}(x, \tilde{u})}{f_X(x) f_{\tilde{U}}(\tilde{u})} \right|_{x=X, \tilde{u}=\tilde{U}} \right]}_{:=B_\tau\text{: Asymptotic Bias}} \end{aligned}$$

Thus, the marginal effect M_τ, consists of an apparent effect A_τ, studied by Firpo et al. (2009), and an asymptotic bias B_τ. Now, we turn our focus to the bias term. Recall that $\tilde{U} := (W,U)$. So we can write

$$B_\tau := \frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\left\{ r(X,W,U) \leq y_\tau \right\} \frac{\partial}{\partial x} \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_{W,U}(w,u)} \bigg|_{x=X,w=W,u=U} \right].$$

where W is actually observable. We actually have

$$\frac{\partial}{\partial x} \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_{W,U}(w,u)} = \frac{\partial}{\partial x} \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_W(w) f_U(u)}$$

Let us define

$$\mathscr{I}(x,w,u) := \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_W(w) f_U(u)}$$

and

$$\mathscr{I}_x(x,w,u) := \frac{\partial}{\partial x} \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_W(w) f_U(u)}.$$

We note that $E\left[\mathscr{I}(X,\tilde{U})\right]$ is called the *mutual information* between X and $\tilde{U}$. Thus, we can write the bias as

$$B_\tau := \frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\left\{ r(X,W,U) \leq y_\tau \right\} \mathscr{I}_x(X,W,U) \right].$$

Remark 1.8. *Unless f_Y is bounded away from 0, it is not possible to bound the bias uniformly across quantiles. If, on the other hand, $f_Y(y) > \gamma$ for some γ and every $y \in \mathscr{Y}$. Then*

$$|B_\tau| \leq \frac{1}{\gamma} E\left[|\mathscr{I}_x(X,W,U)| \right].$$

uniformly across τ.

Remark 1.9. *The sign of the bias B_τ is unlikely to be the same across quantiles. This is a source of difficulty for practitioners when trying to asses the direction bias a priori.*

1.3.1 Omitted Variable Bias

A common source of bias that worries practitioners is omitted variable bias. Suppose that the model is $Y = r(X,W,U)$, but we omit W. The apparent part of the marginal effect, A_τ, remains unchanged and is unaffected by this omission. The following lemma contains a result on the asymptotic bias B_τ.

Theorem 1.8. *Under Assumptions 1.7, 1.8, 1.9, 1.10, 1.11, and either 1) $X,W \perp U$, or 2) $X \perp U|W$, then*

$$B_\tau = -\frac{1}{f_Y(y_\tau)} E\left[\frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} \bigg|_{x=X,w=W} \right] - A_\tau.$$

Thus, the marginal effect M_τ is

$$M_\tau = -\frac{1}{f_Y(y_\tau)} E\left[\frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} \bigg|_{x=X,w=W} \right].$$

This result highlights the correct way to include covariates in the (conditionally) exogenous case. Furthermore, if we additionally require that X and W be independent, then

$$\begin{aligned} M_\tau &= -\frac{1}{f_Y(y_\tau)} E\left[\frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} \bigg|_{x=X,w=W} \right] \\ &= -\frac{1}{f_Y(y_\tau)} E\left[\frac{\partial F_{Y|X=x}(y_\tau)}{\partial x} \bigg|_{x=X} \right] \\ &= A_\tau \end{aligned}$$

and the estimation procedure is much simpler in this case, since we can safely ignore W.

1.4 Conclusion

In this chaper we study the unconditional quantile effect with either a binary treatment or a continuous covariate. In the first case, we find that the unconditional quantile regression estimator that neglects endogeneity can be severely biased. Furthermore, the bias may not be uniform across quantiles, and any attempt to sign the bias *a priori* requires very strong assumptions on the data generating process. More intriguingly, the unconditional quantile regression estimator can be inconsistent even if the treatment status is exogenously determined. This happens when the adjustment given by the derivative of the propensity score is neglected. For example, this is the case when treatment selection is partly determined by covariates that also influence the outcome variable. For the second case, we show how to explicitly obtain a bias formula for the case of the manipulation of a continuous covariate. The bias is also hard to sign a priori, and unlikely to be uniform across quantiles.

Chapter 1, in part, is being prepared for submission for publication, and is coauthored with Yixiao Sun. The dissertation author, Julián Martínez-Iriarte, was the primary author of this chapter.

Chapter 2

Identification and Estimation with an Instrumental Variable

2.1 Introduction

In this chapter we show that if an instrumental variable is available, the UQE can be point identified. We follow Heckman and Vytlacil (1999, 2001a, 2005) and use a threshold-crossing model for treatment selection. In this case, an instrumental variable affects the selection, but it is absent from the outcome equation. We show that the threshold-crossing model implies that individuals who are indifferent between taking up the treatment and not taking up the treatment will drive the unconditional quantile effect. We introduce a new class of marginal treatment effects (MTEs) and show that the unconditional quantile effect can be represented as a weighted average of these MTEs. The MTE we introduce is based on the influence function of the quantile functional. It is related to the marginal treatment effect introduced by Bjorklund and Moffitt (1987) and further studied by Heckman (1997) but is also distinctly different. Identification is achieved using the local instrumental variable approach as in Carneiro and Lee (2009).

A second contribution is to show that the unconditional quantile effect and the marginal policy relevant treatment effect (MPRTE) of Carneiro et al. (2010) belong to the same family of parameters. To the best of our knowledge, this was not previously recognized in either literature. This stems from the fact that our method is more general and allows us to estimate the effect on any (well-behaved) functional of the outcome distribution. In this case, we just need to work with the influence function of this functional. Common examples of a general functional include the quantiles, the mean, and the Gini coefficient.

Finally we develop methods of statistical inference on the UQE when the binary treatment is endogeneous. We take a nonparametric approach but allow the propensity score function to be either parametric or nonparametric. We establish the asymptotic distribution of the UQE estimator. This is a formidable task, as the UQE estimator is a four-step estimator, and we have to pin down the estimation error from each step. Perhaps surprisingly, we show that the error from estimating the propensity score function, either parametrically or nonparametrically, does not affect the asymptotic variance of our UQE estimator.

We are not the first to consider unconditional quantile regressions under endogeneity. Kasy (2016) focuses on the ranking of counterfactual policies and, for the case of discrete regressors, allows for endogeneity. However, one key difference from our approach is that the counterfactual policies analyzed in Kasy (2016) are randomly assigned conditional on a covariate vector. In our setting, selection into treatment follows a threshold-crossing model, where we use the exogenous variation of an instrument to obtain different counterfactual scenarios. Our goal is not to rank potential policies, although our method can be used to rank the class of policies, each of which changes a different instrumental variable. For the case of continuous endogenous covariates, Rothe (2010) shows that the control function approach of Imbens and Newey (2009) can be used to achieve identification. Unlike Rothe (2010), we do not make the unconfoundedness assumption here.

Our general treatment of the problem using functionals is closely related to that of Rothe (2012). Rothe (2012) analyzes the effect of an arbitrary unconditional change in the distribution of a target covariate, either continuous or discrete, on some feature of the distribution of the outcome variable. By assuming a form of conditional independence, for the case of continuous target covariates, Rothe (2012) generalizes the approach of Firpo et al. (2009). However, for the case of a discrete treatment, bounds are obtained by assuming that either the highest-ranked or lowest-ranked individuals enter the program under the new policy.

More recently, Zhou and Xie (2019) provide a new form of the marginal treatment effect parameter that conditions on the propensity score instead of on the whole vector of covariates. Their results allow them to obtain an easy representation of the mean effect when individuals are at the margin of indifference. While our results are also driven by individuals at the margin of indifference, we do not use their redefinition of the marginal treatment effect parameter.

This chapter proceeds as follows: Section 2.2 develops the marginal treatment effect approach and discusses identification. Section 2.3 considers the case of a general functional. Section 2.4 formally establishes the link between the unconditional quantile effect and the marginal policy relevant treatment effect. We consider estimation and inference under a parametric propensity score in Section 2.5 and under a nonparametric propensity score in Section 2.6. We revisit the empirical application of Carneiro et al. (2011) and focus on unconditional quantile effects in Section 2.7. Section 2.8 concludes. We relegate all proofs to the appendix.

2.2 Unconditional Quantile Regressions with an Instrument

In Chapter 1, we have shown that the estimator of Firpo et al. (2009) will be asymptotically biased under endogeneity. It is hard to sign the bias, but more importantly, the bias may not be uniform across quantiles, as shown in Figure 1.4. However, if a special covariate is available, the unconditional quantile effect can be point identified and consistently estimated.

We consider the same model as before. We partition the covariates W into two parts:

$$W = (Z, X).$$

We assume that $Z \in \mathbb{R}$ is a special variable that does not enter the potential outcome equations. In addition, we make the following assumptions, taken directly from Heckman and Vytlacil (1999, 2001a, 2005).

Assumption 2.1. ***Relevance and Exogeneity***

(a) $\mu(Z,X)$ *is a non-degenerate random variable conditional on* X.

(b) (U_0,U_1,V) *is independent of* Z *conditional on* X.

Assumption 2.1(a) is a relevance assumption: for any given level of X, the variable Z can induce some variation in D. Assumption 2.1(b) is referred to as an exogeneity assumption. The two assumptions are essentially the conditions for a valid instrumental variable, hence we will refer to Z as the instrumental variable.

Assumption 2.1(b) allows us to write

$$U_D = F_{V|Z,X}(V|Z,X) = F_{V|X}(V|X).$$

Assumption 2.1(b) then implies that (U_0,U_1,U_D) is independent of Z conditional on X. Based on the value $U_D = u$, we define a marginal treatment effect for the τ-quantile, which will be a basic building block for the unconditional quantile effect.

Definition 2.1. *The marginal treatment effect for the* τ*-quantile is defined as*[1]

$$\text{MTE}_\tau(u,x) = E\left[\mathbb{1}\{Y(0)\le y_\tau\} - \mathbb{1}\{Y(1)\le y_\tau\} \mid U_D = u, X = x\right],$$

where y_τ *is the* τ*-quantile of* $Y = DY(1)+(1-D)Y(0)$, *that is,* $\Pr(Y \le y_\tau) = \tau$.

To aid in the understanding of MTE_τ, we compare it to the marginal treatment effect of Heckman and Vytlacil (1999, 2001a, 2005): $\text{MTE}(u,x) := E[Y(1)-Y(0)\,|U_D = u, X = x]$. For

[1]We could also define

$$\begin{aligned}\text{MTE}_\tau(u,x) &= E\left[\frac{\tau-\mathbb{1}\{Y(1)\le y_\tau\}}{f_Y(y_\tau)} - \frac{\tau-\mathbb{1}\{Y(0)\le y_\tau\}}{f_Y(y_\tau)}\middle| U_D = u, X = x\right] \\ &= \frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\{Y(0)\le y_\tau\} - \mathbb{1}\{Y(1)\le y_\tau\} \mid U_D = u, X = x\right],\end{aligned}$$

but we omit the multiplicative factor $\frac{1}{f_Y(y_\tau)}$ for notational simplicity.

a given individual, $Y(1)-Y(0)$ is the (individual-level) treatment effect, so that the MTE(u,x) is the average treatment effect for individuals with characteristics $U_D=u, X=x$.

Define $\Delta(y_\tau) := \mathbb{1}\{Y(0)\le y_\tau\} - \mathbb{1}\{Y(1)\le y_\tau\}$, which is the argument of the conditional expectation in our definition of MTE$_\tau$. The random variable $\Delta(y_\tau)$ can take three values:

$$\Delta(y_\tau) = \begin{cases} 1 \text{ if } Y(0)\le y_\tau \text{ and } Y(1) > y_\tau \\ 0 \text{ if } \left[Y(0) > y_\tau \text{ and } Y(1) > y_\tau\right] \text{ or } \left[Y(0)\le y_\tau \text{ and } Y(1)\le y_\tau\right] \\ -1 \text{ if } Y(0) > y_\tau \text{ and } Y(1)\le y_\tau \end{cases}$$

For a given individual, $\Delta(y_\tau)=1$ when the treatment induces the individual to "cross" the τ-quantile y_τ of Y from below, and $\Delta(y_\tau)=-1$ when the treatment induces the individual to "cross" the τ-quantile y_τ of Y from above. In the first case the individual benefits from the treatment while in the second case the treatment harms her. The intermediate case, $\Delta(y_\tau)=0$, occurs when the treatment induces no quantile crossing of any type. Thus the unconditional expected value $E[\Delta(y_\tau)]$ equals the difference between the proportion of individuals who benefit from the treatment and the proportion of individuals who are harmed by it. For the UQE, whether the treatment is beneficial or harmful is measured in terms of quantile crossing. Among the individuals with characteristics $U_D=u, X=x$, MTE$_\tau(u,x)$ is then the difference between the proportion of individuals who benefit from the treatment and the proportion of individuals who are harmed by it. Thus, MTE$_\tau(u,x)$ is positive if more individuals increase their outcome above y_τ, and it is negative if more individuals decrease their outcome below y_τ.

MTE$_\tau(u,x)$ is different from the quantile analogue of the marginal treatment effect of Carneiro and Lee (2009), which is defined as $F^{-1}_{Y(1)|U_D,X}(\tau|u,x) - F^{-1}_{Y(0)|U_D,X}(\tau|u,x)$. MTE$_\tau(u,x)$ is proportional to the difference of (the conditional expectations of) the influence functions for the τ-quantile of $Y(0)$ and $Y(1)$. The proportionality factor is $f_Y(y_\tau)$, the unconditional density of Y evaluated at the τ-quantile y_τ of Y; see footnote 1.

2.2.1 UQE under Location Shift, Revisited

In the new setting with an instrument, it is worthwhile revisiting Corollary 1.5. Note that if we induce a location shift in the benefit function, the unconditional quantile effect $M_{\tau,\mu}$ is still given by Corollary 1.5 as long as the same assumptions hold for $W=(Z,X)$. Using Assumption 2.1(b), we can obtain a representation of $M_{\tau,\mu}$ in terms of the marginal treatment effect for the τ-quantile. This representation will be useful for identification, as we show later in Section 2.2.3.

Theorem 2.1. *Let Assumptions 1.4–1.6 and Assumption 2.1(b) hold. Assume further that* $f_Y(y_\tau)>0$*. Then under the location shift in the benefit function given by (1.11), we have*

$$M_{\tau,\mu}=\frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}}\text{MTE}_\tau(P(w),x)\dot{P}^{\mu}(w)\,f_W(w)\,dw,$$

where

$$\dot{P}^{\mu}(w)=\frac{f_{V|X}(\mu(w)|x)}{E\left[f_{V|X}(\mu(W)|X)\right]}.$$

The main difference between Theorem 2.1 and Corollary 1.5 is that conditioning on $W=w$ has been replaced by conditioning on $X=x$ only. This is possible because of Assumption 2.1(b). When (U_0,U_1,V) is independent of Z conditional on X, we know that (U_0,U_1) is independent of Z given (U_D,X), as

$$\begin{aligned}f_{U,Z|U_D,X}(u,z|u_D,x) &= \frac{f_{U,U_D,Z|X}(u,u_D,z|x)}{f_{U_D|X}(u_D|x)}=\frac{f_{U,U_D|X}(u,u_D|x)}{f_{U_D|X}(u_D|x)}\cdot f_{Z|X}(z|x)\\ &= f_{U|U_D,X}(u|u_D,x)\cdot f_{Z|U_D,X}(z|u_D,x).\end{aligned} \tag{2.1}$$

Given this and the assumption that Z does not enter $Y(0)$ or $Y(1)$, we have, for $w=(z,x)$,

$$\begin{aligned}&E\left[\mathbb{1}\{Y(0)\le y_\tau\}-\mathbb{1}\{Y(1)\le y_\tau\}\mid U_D=u,W=w\right]\\ =\;& E\left[\mathbb{1}\{Y(0)\le y_\tau\}-\mathbb{1}\{Y(1)\le y_\tau\}\mid U_D=u,Z=z,X=x\right]\\ =\;& E\left[\mathbb{1}\{Y(0)\le y_\tau\}-\mathbb{1}\{Y(1)\le y_\tau\}\mid U_D=u,X=x\right].\end{aligned}$$

Thus we can indeed replace "conditioning on $W = w$" by "conditioning on $X = x$."

2.2.2 UQE under Instrument Intervention

In this subsection, we consider a different kind of manipulation of the propensity score: we shift the location of the instrumental variable Z. To be precise, suppose we shift Z to $Z_\delta = Z + g(W)s_z(\delta)$, where $g(\cdot)$ is a measurable function to be determined in the sense that it is user specified. Recall that we partition the covariates as $W = (Z, X)$. Note that while $s_z(\delta)$ is the same for all individuals, $g(W)$ depends on the value of W and hence it is individual specific. Thus, we allow the intervention to be heterogeneous.

A simple homogenous intervention is obtained by setting $g(\cdot) \equiv 1$, in which case we have an additive shift in Z. An example of an heterogeneous intervention is given by $g(W) = Z$, in which case we obtain a multiplicative shift of order $1 + s_z(\delta)$. Both of these cases have been studied by Carneiro et al. (2010).

Selection into treatment is now governed by

$$D_\delta = \mathbb{1}\{V \leq \mu(Z + g(W)s_z(\delta), X)\}, \tag{2.2}$$

where, for a given $g(\cdot)$, the choice of $s_z(\delta)$ guarantees that $\Pr(D_\delta = 1) = p + \delta$.

Lemma 2.1. *Assume that (i) (V, X) are absolutely continuous random variables with joint density $f_{V,X}(v,x)$ given by $f_{V|X}(v|x) f_X(x)$; (ii) $f_{V|X}(v|x)$ is continuous in v for almost all $x \in \mathscr{X}$; (iii) $\mu(z,x)$ is continuously differentiable in z for almost all $x \in \mathscr{X}$; (iv) letting $\mu_z'(z,x)$ be the partial derivative of $\mu(z,x)$ with respect to z, we have*

$$E \sup_{\delta \in N_\varepsilon} \left[f_{V|X}(\mu(Z + g(W)s_z(\delta), X)|X) \mu_z'(Z + g(W)s_z(\delta), X) \right] g(W) < \infty;$$

(v) for each $\delta \in N_\varepsilon$,

$$E\left[f_{V|X}\left(\mu\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)|X\right)\mu_z'\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)g(W)\right] \neq 0.$$

Then

$$\begin{aligned}\left.\frac{\partial s_z(\delta)}{\partial \delta}\right|_{\delta=0} &= \frac{1}{E\left[f_{V|W}\left(\mu\left(W\right)|W\right)\mu_z'\left(W\right)g(W)\right]},\\ \left.\frac{\partial P_\delta(z,x)}{\partial \delta}\right|_{\delta=0} &= \frac{f_{V|W}\left(\mu(w)|w\right)\mu_z'\left(w\right)g(w)}{E\left[f_{V|W}\left(\mu\left(W\right)|W\right)\mu_z'\left(W\right)g(W)\right]}.\end{aligned}$$

Theorem 2.2. *Let Assumptions 1.4–1.6 and 2.1(b), and the assumptions of Lemma 2.1 hold. Assume further that* $f_Y(y_\tau) > 0$*. Then, the unconditional quantile effect of the shift in Z given in (2.2) is*

$$M_{\tau,z} = \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \mathrm{MTE}_\tau\left(P\left(w\right),x\right)\dot{P}^z\left(w\right)f_W\left(w\right)dw,$$

where

$$\dot{P}^z(w) = \left.\frac{\partial P_\delta(w)}{\partial \delta}\right|_{\delta=0} = \frac{f_{V|X}\left(\mu(w)|x\right)\mu_z'\left(w\right)g(w)}{E\left[f_{V|X}\left(\mu\left(W\right)|X\right)\mu_z'\left(W\right)g\left(W\right)\right]}.$$

It can be seen that the main difference between $M_{\tau,\mu}$ in Theorem 2.1 and $M_{\tau,z}$ in Theorem 2.2 lies in the adjustment given by the derivative of the modified propensity score $P_\delta(w)$. For $M_{\tau,z}$, the adjustment to the population weights includes the derivative of the benefit function, $\mu_z'(\cdot)$, and the function $g(\cdot)$. Different adjustments lead to different distributions of W over the marginal subpopulation.

Note that if $\mu_z'(w)$ is known and different from 0 for all $w \in \mathscr{W}$, then choosing $g(w) = 1/\mu_z'(w)$ yields $M_{\tau,z} = M_{\tau,\mu}$. In the special case where $g(w) = \mu_z'(w) = 1$ for all $w \in \mathscr{W}$, the two effects coincide. Note that the condition $\mu_z'(w) = 1$ for all $w \in \mathscr{W}$ holds only if

$$\mu\left(z,x\right) = z + \tilde{\mu}\left(x\right)$$

for some function $\tilde{\mu}(\cdot)$. So, if we think of $\mu(\cdot,\cdot)$ as the utility function, then the utility function is required to take a quasilinear form with z as the numeraire. Shifts to the benefit function and shifts to the instrument (upon choosing $g(\cdot) \equiv 1$) are then equivalent.

2.2.3 Identification of MTE$_\tau$

To investigate the identifiability of $M_{\tau,z}$, we study the identifiability of MTE$_\tau$ and the weight function or the RN derivative $\left.\frac{\partial P_\delta(w)}{\partial \delta}\right|_{\delta=0}$ separately. The proposition below shows that MTE$_\tau(u,x)$ is identified for every $u = P(w)$ for some $w \in \mathscr{W}$.

Proposition 2.1. *Let Assumptions 1.5(a), 1.5(b), and 2.1(b) hold. Then, for every $u = P(w)$ with $w \in \mathscr{W}$, we have*

$$\text{MTE}_\tau(u,x) = -\frac{\partial E\left[\mathbb{1}\{Y \le y_\tau\}\,|P(W) = u, X = x\right]}{\partial u}. \tag{2.3}$$

Proposition 2.1 can be proved using Theorem 1 in Carneiro and Lee (2009). In the supplementary appendix, we provide a self-contained proof that is directly connected to the idea of shifting the propensity score.

The key results that we use to establish Proposition 2.1 are

$$\begin{aligned} E\left[\mathscr{G}(Y(1))\,|P(W) = P(w), U_D \le P(w), X = x\right] &= E\left[\mathscr{G}(Y(1))|U_D \le P(w), X = x\right], \\ E\left[\mathscr{G}(Y(0))|P(W) = P(w), U_D > P(w), X = x\right] &= E\left[\mathscr{G}(Y(0))|U_D > P(w), X = x\right], \end{aligned}$$

where $\mathscr{G}(\cdot)$ is a bounded function with $\mathscr{G}(\cdot) = \mathbb{1}\{\cdot \le y_\tau\}$ as a special case. These results hold under the assumption of instrument exogeneity. Without them, we have only that

$$\begin{aligned} \frac{\partial E\left[\mathscr{G}(Y)|P(W) = u, X = x\right]}{\partial u} &= E\left[\mathscr{G}(Y(1))|P(W) = u, U_D = u, X = x\right] \\ &- E\left[\mathscr{G}(Y(0))|P(W) = u, U_D = u, X = x\right] \\ &+ \int_0^u \frac{\partial E\left[\mathscr{G}(Y(1))|P(W) = u, U_D = \tilde{u}, X = x\right]}{\partial u} d\tilde{u} \\ &+ \int_u^1 \frac{\partial E\left[\mathscr{G}(Y(0))|P(W) = u, U_D = \tilde{u}, X = x\right]}{\partial u} d\tilde{u}. \end{aligned} \tag{2.4}$$

Under the assumption of instrument exogeneity (and the assumption that Z does not affect the potential outcomes directly), we have that

$$\frac{\partial E\left[\mathscr{G}(Y(1))|P(W)=u,U_D=\tilde{u},X=x\right]}{\partial u}=0$$

when $\tilde{u}\leq u$, and that

$$\frac{\partial E\left[\mathscr{G}(Y(0))|P(W)=u,U_D=\tilde{u},X=x\right]}{\partial u}=0$$

when $\tilde{u}>u$. Hence, the last two terms in (2.4) disappear, and conditioning on $P(W)=u$ in the first two terms can be dropped. Therefore, a key identification assumption for MTE_τ is the assumption of instrument exogeneity.

2.2.4 Identification of the RN Derivative

In this subsection, we investigate the identification of the RN derivative $\dot{P}^z(w)$ in the representation of $M_{\tau,z}$. By Lemma 2.1 together with Assumption 2.1(b), we have

$$\dot{P}^z(w)=\left.\frac{\partial P_\delta(w)}{\partial\delta}\right|_{\delta=0}=\frac{f_{V|X}\left(\mu(w)|x\right)\mu_z'(w)\,g(w)}{E\left[f_{V|X}\left(\mu\left(W\right)|X\right)\mu_z'\left(W\right)g\left(W\right)\right]},$$

where $\mu_z'(w):=\frac{\partial\mu(z,x)}{\partial z}$. Under Assumption 2.1(b), the propensity score becomes

$$\begin{aligned}P(w)&=\Pr\left(V\leq\mu\left(W\right)|W=w\right)\\&=F_{V|Z,X}(\mu(z,x)|z,x)=F_{V|X}(\mu(z,x)|x).\end{aligned}\qquad(2.5)$$

Therefore,

$$\frac{\partial P(w)}{\partial z}=f_{V|X}(\mu(z,x)|x)\mu_z'(z,x)=f_{V|X}(\mu(w)|x)\mu_z'(w).$$

It is now clear that $\dot{P}^z(w)$ can be represented using $\frac{\partial P(w)}{\partial z}$ and $g(w)$. We formalize this in the following proposition.

Proposition 2.2. *Let Assumption 2.1(b) and the assumptions in Lemma 2.1 hold. Then*

$$\dot{P}^z(w) = \left.\frac{\partial P_\delta(w)}{\partial \delta}\right|_{\delta=0} = \frac{\frac{\partial P(w)}{\partial z} g(w)}{E\left[\frac{\partial P(W)}{\partial z} g(W)\right]}. \tag{2.6}$$

Since $g(w)$ is known and $\frac{\partial P(w)}{\partial z}$ is identified, $\dot{P}^z(w)$ is also identified. As in the case of MTE_τ identification, Assumption 2.1(b) plays a key role in identifying $\dot{P}^z(w)$. Without the assumption that V is independent of Z conditional on X, we can have only that

$$\dot{P}^z(w) = \frac{f_{V|W}(\mu(w)|w)\,\mu_z'(w)\,g(w)}{E\left[f_{V|W}(\mu(W)|W)\,\mu_z'(W)\,g(W)\right]}$$

and

$$\frac{\partial P(w)}{\partial z} = f_{V|Z,X}(\mu(w)|w)\mu_z'(w) + \frac{\partial F_{V|Z,X}(\mu(w)|\tilde{z},x)}{\partial \tilde{z}}|_{\tilde{z}=z}.$$

The presence of the second term in the above equation invalidates the identification result in (2.6).

Using Propositions 2.1 and 2.2, we can represent $M_{\tau,z}$ as

$$\begin{aligned} M_{\tau,z} = &-\frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \frac{\partial E\left[\mathbb{1}\{Y \le y_\tau\}\,|P(W) = P(w), X = x\right]}{\partial P(w)} \\ &\times \frac{\frac{\partial P(w)}{\partial z} g(w)}{E\left[\frac{\partial P(W)}{\partial z} g(W)\right]} f_W(w)\,dw. \end{aligned} \tag{2.7}$$

All objects in the above are point identified, hence $M_{\tau,z}$ is point identified.

We note that, in general, $M_{\tau,\mu}$ is not point identified. Even if there is a valid instrument such that $\text{MTE}_\tau(u,x)$ is identified, $\dot{P}^\mu(w)$, the RN derivative in the definition of $M_{\tau,\mu}$, may not be point identified. Observing that

$$\dot{P}^{\mu}(w) = \left.\frac{\partial P_{\delta}(w)}{\partial \delta}\right|_{\delta=0} = \frac{f_{V|X}(\mu(w)|x)}{E\left[f_{V|X}(\mu(W)|X)\right]},$$

we see that, in general, $\dot{P}^{\mu}(w)$ can be identified using the instrument in only the special case where $\mu_z'(w) = 1$ for all $w \in \mathscr{W}$. In this special case, $\dot{P}^{\mu}(w) = \dot{P}^{z}(w)$.

When $\dot{P}^{\mu}(w)$ is not identified, $M_{\tau,\mu}$ is also not identified. The most we can do in such a case is to bound the unconditional quantile effect. Suppose $\dot{P}^{\mu}(w)$ has the same sign so that $\dot{P}^{\mu}(w) \geq 0$ for all $w \in \mathscr{W}$.[2] Then we have

$$\begin{aligned} M_{\tau,\mu} &= \int_{\mathscr{W}} \text{MTE}_{\tau}(P(w),x)\dot{P}^{\mu}(w)f_W(w)dw \\ &\in \left[\inf_{w\in\mathscr{W}} \text{MTE}_{\tau}(P(w),x), \sup_{w\in\mathscr{W}} \text{MTE}_{\tau}(P(w),x)\right] \end{aligned}$$

because $\int_{\mathscr{W}} \dot{P}^{\mu}(w)f_W(w) = 1$. We leave the details of the bound approach under partial identification for future research.

2.3 General Unconditional Policy Effects under Endogeneity

Let $\mathscr{F}^*$ be the space of finite signed measures ν on $\mathscr{Y} \subseteq \mathbb{R}$ with distribution function $F_{\nu}(y) = \nu(-\infty, y]$ for $y \in \mathscr{Y}$. We endow $\mathscr{F}^*$ with the usual supremum norm: for two distribution functions F_{ν_1} and F_{ν_2} associated with the signed measures ν_1 and ν_2 on $\mathscr{Y}$, we define $\|F_{\nu_1} - F_{\nu_2}\| = \sup_{y\in\mathscr{Y}} |F_{\nu_1}(y) - F_{\nu_2}(y)|$. In this section, we consider a general functional $\rho : \mathscr{F}^* \to \mathbb{R}$ and study the general unconditional effect.

The baseline model is the same as in Section 2.2. As before, we modify the propensity score to improve the treatment take-up rate from p to $p+\delta$. The (general) unconditional policy effect is defined to be the marginal change in $\rho(F_{Y_{\delta}})$ in the limit as δ goes to 0.

[2]Given that $\int_{\mathscr{W}} \dot{P}^{\mu}(w)f_W(w) = 1$, it is impossible that $\dot{P}^{\mu}(w) \leq 0$ for all $w \in \mathscr{W}$.

Definition 2.2. ***General Unconditional Policy Effect***

The general unconditional policy effect for the functional ρ is defined as

$$M_\rho = \lim_{\delta \to 0} \frac{\rho(F_{Y_\delta}) - \rho(F_Y)}{\delta}$$

whenever this limit exists.

The definition above is the same as that of the marginal partial distributional policy effect defined in Rothe (2012). The term we use is closer to those in Firpo et al. (2009).

2.3.1 Characterization of the General Unconditional Policy Effect

We first consider a Hadamard differentiable functional ρ. For completeness, we provide the definition of Hadamard differentiability below.

Definition 2.3. $\rho : \mathscr{F}^* \to \mathbb{R}$ *is Hadamard differentiable at $F \in \mathscr{F}^*$ if there exists a linear and continuous functional $\dot{\rho}_F : \mathscr{F}^* \to \mathbb{R}$ such that for any $G \in \mathscr{F}^*$ and $G_\delta \in \mathscr{F}^*$ with*

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} |G_\delta(y) - G(y)| = 0,$$

we have

$$\lim_{\delta \to 0} \frac{\rho(F_Y + \delta G_\delta) - \rho(F_Y)}{\delta} = \dot{\rho}_F(G).$$

Recall that by Lemma 1.3 we have the expansion

$$\begin{aligned} F_{Y_\delta}(y) &= F_Y(y) \\ &+ \delta E\left[\left\{F_{Y(1)|U_D,W}(y|P(W),W) - F_{Y(0)|U_D,W}(y|P(W),W)\right\}\dot{P}(W)\right] + R_F(\delta;y), \end{aligned}$$

where $\sup_{y \in \mathscr{Y}} |R_F(\delta;y)| = o(|\delta|)$ as $\delta \to 0$. Taking

$$G_\delta(y) = \frac{1}{\delta}\left[F_{Y_\delta}(y) - F_Y(y)\right]$$

and

$$G(y) = E\left[\left\{F_{Y(1)|U_D,W}(y|P(W),W) - F_{Y(0)|U_D,W}(y|P(W),W)\right\}\dot{P}(W)\right], \tag{2.8}$$

we have

$$\lim_{\delta\to 0}\sup_{y\in\mathscr{Y}}|G_\delta(y) - G(y)| = 0.$$

Hence, under the Hadamard differentiability of ρ, we obtain

$$\begin{aligned} M_\rho &= \lim_{\delta\to 0}\frac{\rho(F_{Y_\delta}) - \rho(F_Y)}{\delta} = \lim_{\delta\to 0}\frac{\rho(F_Y + \delta G_\delta) - \rho(F_Y)}{\delta} = \dot{\rho}_F(G) \\ &= \int_{\mathscr{Y}}\psi(y,\rho,F_Y)dG(y), \end{aligned}$$

where $\psi(y,\rho,F_Y)$ is the influence function of ρ at F_Y. Plugging (2.8) into this result yields the following theorem.

Theorem 2.3. *Let the assumptions in Lemma 1.3 (i.e., Assumptions 1.4–1.6) hold. Assume further that $\rho : \mathscr{F}^* \to \mathbb{R}$ is Hadamard differentiable. Then*

$$M_\rho = \int_{\mathscr{Y}}\psi(y,\rho,F_Y)E\left[\left\{f_{Y(1)|U_D,W}(y|P(W),W) - f_{Y(0)|U_D,W}(y|P(W),W)\right\}\dot{P}(W)\right]dy. \tag{2.9}$$

Define

$$\mathrm{MTE}_\rho(u,w) = E\left[\psi(Y(1),\rho,F_Y) - \psi(Y(0),\rho,F_Y)|U_D = u, W = w\right]. \tag{2.10}$$

Then

$$M_\rho = \int_{\mathscr{W}}\mathrm{MTE}_\rho(u,w)\dot{P}(w)f_W(w)dw. \tag{2.11}$$

Hence, the general unconditional policy effect M_ρ can be represented as a weighted average of $\mathrm{MTE}_\rho(u,w)$ over the marginal subpopulation.

Consider the quantile functional: $\rho_\tau(F) = F^{-1}(\tau)$. It is well known that this functional

$\rho_\tau(\cdot)$ is Hadamard differentiable, and its influence function is

$$\psi(y, \rho_\tau, F_Y) = \frac{\tau - \mathbb{1}\{y \le y_\tau\}}{f_Y(y_\tau)}. \tag{2.12}$$

Plugging this into (2.10) yields

$$\text{MTE}_\tau(u, w) = \frac{1}{f_Y(y_\tau)} E\left[\mathbb{1}\{Y(0) \le y_\tau\} - \mathbb{1}\{Y(1) \le y_\tau\} \mid U_D = u, W = w\right].$$

This is the same as MTE_τ in Definition 2.1 except for the scaling factor $f_Y(y_\tau)$ and the absence of the instrument in the conditioning set. The representation in Theorem 2.3 is then exactly the same as the representation in Theorem 1.3.

Following Firpo et al. (2009), we may construct an unconditional regression using $\psi(y, \rho_\tau, F_Y)$ as the dependent variable and D and other covariates as the independent variables. Like the UQR estimator, such an estimator will be inconsistent for M_ρ, and its asymptotic bias can be similarly decomposed into two sources. The identification of M_ρ under instrument intervention for a general ρ can be established in the same way as that for the quantile functional. We omit the details here.

While Theorem 2.3 covers general functionals, it does not cover the mean functional $\rho(F) = \int_{\mathscr{Y}} y dF(y)$ unless $\mathscr{Y}$ is a bounded set. When $\mathscr{Y}$ is unbounded, the mean functional is not continuous on $(\mathscr{F}^*, \|\cdot\|_\infty)$ and hence is not Hadamard differentiable (see Exercise 7 in Chapter 20 in van der Vaart (1998)). In such a case, we opt for a direct approach by showing that

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{\mathscr{Y}} y dR_F(\delta; y) = 0, \tag{2.13}$$

so that

$$\begin{aligned}
M_\rho &= \int_{\mathscr{Y}} yE\left[\left\{f_{Y(1)|U_D,W}(y|P(W),W) - f_{Y(0)|U_D,W}(y|P(W),W)\right\}\dot{P}(W)\right]dy \\
&+ \lim_{\delta\to 0}\frac{1}{\delta}\int_{\mathscr{Y}} ydR_F(\delta;y) \\
&= \int_{\mathscr{Y}} yE\left[\left\{f_{Y(1)|U_D,W}(y|P(W),W) - f_{Y(0)|U_D,W}(y|P(W),W)\right\}\dot{P}(W)\right]dy.
\end{aligned}$$

The result in (2.13) holds if the following stronger version of Assumption 1.6 holds:

Assumption 2.2. ***Stronger Domination Conditions*** *For* $d = 0, 1$

$$\begin{aligned}
\int_{\mathscr{Y}(d)} \sup_{\delta\in N_\varepsilon} y\cdot f_{Y(d)|D_\delta}(y|d)dy &< \infty, \\
\int_{\mathscr{Y}(d)} \sup_{\delta\in N_\varepsilon}\left| y\cdot\frac{\partial f_{Y(d)|D_\delta}(y|d)}{\partial\delta}\right| dy &< \infty.
\end{aligned}$$

Corollary 2.4. *Let Assumptions 1.4, 1.5, and 2.2 hold. Then for the mean functional, we have*

$$M_\rho = E\left\{[Y(1) - Y(0)\,|U_D = P(W),W]\,\dot{P}(W)\right\}. \tag{2.14}$$

2.4 Relationship to Marginal Policy Relevant Treatment Effects

In this section we take a closer look at the relationship between general unconditional effects under endogeneity and marginal policy relevant treatment effects. A policy relevant treatment effect is a comparison of two policies with different incentives for participation. It is assumed that the potential outcomes remain the same. Using the notation of the previous sections, consider a baseline policy where $\delta = 0$, versus an alternative policy where $\delta > 0$. As before, $p = \Pr(D = 1)$, whereas $p + \delta = \Pr(D_\delta = 1)$. Here, D_δ is the treatment status under the alternative policy. Suppose we are interested in the effect of this policy change on the unconditional mean of the observed outcomes. To this end, Heckman and Vytlacil (2001b, 2005) consider the *policy relevant treatment effect* defined as

$$\text{PRTE}_\delta = \frac{E(Y_\delta) - E(Y)}{E(D_\delta) - E(D)} = \frac{E(Y_\delta) - E(Y)}{\delta}. \tag{2.15}$$

Taking the limit $\delta \to 0$ yields the *marginal policy relevant treatment effect* (MPRTE) of Carneiro et al. (2010):

$$\text{MPRTE} = \lim_{\delta \to 0} \text{PRTE}_\delta.$$

Following Heckman and Vytlacil (2001b, 2005), we can show that MPRTE can be represented in terms of the following marginal treatment effect:

$$\text{MTE}(u,x) := E\left[Y(1) - Y(0) | U_D = u, X = x\right].$$

We note that $\text{MTE}(u,x)$ is different from $\text{MTE}_\tau(u,x)$ in Definition 2.1. To simplify the notation, we drop the covariate X. Let $\mathscr{P}$ and $\mathscr{P}_\delta$ be the propensity scores, that is, $\mathscr{P} = \Pr(D = 1|Z)$ and $\mathscr{P}_\delta = \Pr(D_\delta = 1|Z)$. This new notation for the propensity scores, $\mathscr{P}$ and $\mathscr{P}_\delta$, suppresses their dependence on Z and highlight that $\mathscr{P}$ and $\mathscr{P}_\delta$ are themselves random variables. Let $f_{\mathscr{P}_\delta}(\cdot)$ and $F_{\mathscr{P}_\delta}(\cdot)$ be the pdf and CDF of $\mathscr{P}_\delta$, respectively. When $\delta = 0$, we denote the pdf and CDF of $\mathscr{P}$ by $f_{\mathscr{P}}(\cdot)$ and $F_{\mathscr{P}}(\cdot)$, respectively. Then

$$\text{MPRTE} = -\int_0^1 \text{MTE}(u) \frac{\partial F_{\mathscr{P}_\delta}(u)}{\partial \delta}\bigg|_{\delta=0} du. \tag{2.16}$$

A proof is given in the supplementary appendix.

To obtain an expression for $\frac{\partial F_{\mathscr{P}_\delta}(u)}{\partial \delta}|_{\delta=0}$, we consider the special case $\mu(Z) = \gamma Z$ and $\gamma > 0$. This form of $\mu(\cdot)$ is a simplified version of Assumption B-1 in Carneiro et al. (2010). In this case, $\mathscr{P} = F_V(\gamma Z)$ and $\mathscr{P}_\delta = F_V(\gamma Z + \gamma s_z(\delta)$. Consider a constant shift of magnitude $s_z(\delta)$

in Z (i.e., $g(\cdot) = 1$). Under this shift, the participation rate increases by δ so that

$$E(\mathscr{P}_\delta) = \Pr(D_\delta = 1) = \Pr(D = 1) + \delta = E(\mathscr{P}) + \delta.$$

We have

$$\begin{aligned} F_{\mathscr{P}_\delta}(u) &= \Pr(\mathscr{P}_\delta \leq u) = \Pr(F_V(\gamma Z + \gamma s_z(\delta)) \leq u) \\ &= \Pr(\gamma Z + \gamma s_z(\delta) \leq F_V^{-1}(u)) = \Pr(F_V(\gamma Z) \leq F_V(F_V^{-1}(u) - \gamma s_z(\delta))) \\ &= F_{\mathscr{P}}(F_V(F_V^{-1}(u) - \gamma s_z(\delta))). \end{aligned} \tag{2.17}$$

Differentiating (2.17) with respect to δ, we get

$$\frac{\partial F_{\mathscr{P}_\delta}(u)}{\partial \delta} = -f_{\mathscr{P}_\delta}(F_V(F_V^{-1}(u) - \gamma s_z(\delta)))f_V(F_V^{-1}(u) - \gamma s_z(\delta))\gamma \frac{\partial s_z(\delta)}{\partial \delta}. \tag{2.18}$$

Using Lemma 2.1 and setting $g(\cdot) = 1$, we have

$$\left. \frac{\partial s_z(\delta)}{\partial \delta} \right|_{\delta=0} = \frac{1}{\int_{\mathscr{Z}} f_V(\gamma z)\gamma f_Z(z)dz}. \tag{2.19}$$

Evaluating (2.18) at $\delta = 0$ and plugging in (2.19), we get

$$\left. \frac{\partial F_{\mathscr{P}_\delta}(u)}{\partial \delta} \right|_{\delta=0} = -\frac{f_{\mathscr{P}}(u)f_V(F_V^{-1}(u))}{\int_{\mathscr{Z}} f_V(\gamma z)f_Z(z)dz}. \tag{2.20}$$

Now, the marginal policy relevant treatment effect is

$$\text{MPRTE} = \int_0^1 \text{MTE}(u)\frac{f_{\mathscr{P}}(u)f_V(F_V^{-1}(u))}{\int_{\mathscr{Z}} f_V(\gamma z)f_Z(z)dz}du.$$

Consider the change of variable $u = P(z)$, where $P(z) = F_V(\gamma z)$ in this particular case.

Then $du = \gamma f_V(\gamma z)dz$ and $F_V^{-1}(u) = \gamma z$. Note also that

$$F_{\mathscr{P}}(u) = \Pr(\mathscr{P} \le u) = \Pr(F_V(\gamma Z) \le u) = F_Z(\gamma^{-1} F_V^{-1}(u)).$$

Therefore, the density $f_{\mathscr{P}}(u)$ is

$$f_{\mathscr{P}}(u) = f_Z(\gamma^{-1} F_V^{-1}(u)) \frac{\gamma^{-1}}{f_V(F_V^{-1}(u))} = f_Z(z) \frac{\gamma^{-1}}{f_V(F_V^{-1}(u))},$$

and the MPRTE becomes

$$\text{MPRTE} = \int_{\mathscr{Z}} \text{MTE}(P(z)) \frac{f_V(\gamma z)}{\int_{\mathscr{Z}} f_V(\gamma z) f_Z(z) dz} f_Z(z) dz = \int_{\mathscr{Z}} \text{MTE}(P(z)) \dot{P}(z) f_Z(z) dz.$$

This result is precisely the one in (2.14) in Corollary 2.4 after dropping the covariate X. This is also Example 2 in Carneiro et al. (2010) for the case where, in their notation, $q_\alpha(t) = t + \alpha$.

A formal proof of the equivalence of MPRTE to M_ρ when ρ is the mean functional (as in Corollary 2.4) in more general cases is sketched in the appendix. While our results cover MPRTE as a special case, they also cover more general unconditional policy effects.

2.5 Estimation and Inference under a Parametric Propensity Score

This section is devoted to the estimation and inference of the UQE under instrument intervention, as described in Section 2.2. We assume that the propensity score function is parametric, and we leave the case with a nonparametric propensity score to the next section.

Letting

$$m_0(y_\tau, P(w), x) := E\left[\mathbb{1}\left\{Y \le y_\tau\right\} | P(W) = P(w), X = x\right]$$

and using (2.7), we have

$$
\begin{aligned}
M_{\tau,z} &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{W}} \mathrm{MTE}_\tau \left(P\left(w\right), x\right) \dot{P}^z\left(w\right) f_W\left(w\right) dw \\
&= -\frac{1}{f_Y(y_\tau)} \left\{ E\left[\frac{\partial P(W)}{\partial z} g\left(W\right) \right] \right\}^{-1} \int_{\mathscr{W}} \frac{\partial m_0(y_\tau, P(w), x)}{\partial z} g(w) f_W\left(w\right) dw.
\end{aligned}
$$

Note that the presence of $g\left(\cdot\right)$ amounts to a change of measure (from a measure with density $f_W(w)$ to a measure with density $g(w)f_W(w)$). In order to simplify the notation, we set $g(\cdot) \equiv 1$ for the remainder of this chapter. The case with a general but known $g(\cdot)$ can be handled with straightforward modifications.

With $g(\cdot) = 1$, we write the parameter $M_{\tau,z}$ as

$$
M_{\tau,z} = -\frac{1}{f_Y(y_\tau)} E\left[\frac{\partial P(W)}{\partial z} \right]^{-1} E\left[\frac{\partial m_0(y_\tau, P(W), X)}{\partial z} \right]. \tag{2.21}
$$

$M_{\tau,z}$ consists of two average derivatives and a density evaluated at a point. As shown by Newey (1994), the two average derivatives are $\sqrt{n}$-estimable. However, the density at a point cannot be estimated at the usual $\sqrt{n}$ rate unless a parametric model is imposed.

First, we make an assumption concerning the dimensions of the variables.

Assumption 2.3. *The covariate vector X is an element of $\mathbb{R}^{d_X}$ and $Z \in \mathbb{R}$.*

In empirical applications, we may have a few exogenous variables that affect the treatment choice but not the outcome of interest directly. The instrument Z can be any one of these exogenous variables, and the rest of the exogenous variables become part of the covariate vector X. The unconditional effect is specific to the instrumental variable Z that we choose to intervene in order to improve the treatment adoption rate.

The rest of this section is structured as follows: in Section 2.5.1 we establish the rate of convergence of the two-step estimator of $f_Y(y_\tau)$; in Section 2.5.2 we find the asymptotic distribution of the terms associated with the propensity score; and in Section 2.5.3 we establish the asymptotic distribution of $\hat{M}_{\tau,z}$ and construct a pivotal test statistic for testing the null of a zero effect.

For a given sample $\{O_i = (Y_i, Z_i, X_i, D_i)\}_{i=1}^n$, we use $\mathbb{P}_n$ to denote the empirical measure. The expectation of a function $\chi(O)$ with respect to $\mathbb{P}_n$ is $\mathbb{P}_n\chi = n^{-1}\sum_{i=1}^n \chi(O_i)$.

2.5.1 Two-Step Density Estimator

For a given τ, we estimate y_τ using the (generalized) inverse of the empirical distribution function of Y:

$$\hat{y}_\tau = \inf\{y : \mathbb{F}_n(y) \geq \tau\},$$

where

$$\mathbb{F}_n(y) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{Y_i \leq y\}.$$

The following asymptotic result can be found in Serfling (1980).[3]

Lemma 2.2. *If the density $f_Y(\cdot)$ of Y is positive and continuous at y_τ, then*

$$\hat{y}_\tau - y_\tau = \mathbb{P}_n\psi_Q(y_\tau) + o_p(n^{-1/2}),$$

where

$$\psi_Q(y_\tau) := \frac{\tau - \mathbb{1}\{Y \leq y_\tau\}}{f_Y(y_\tau)}.$$

We use a kernel density estimator to estimate $f_Y(y)$. We maintain the following assumption on the kernel function.

Assumption 2.4. ***Kernel Assumption***

The kernel function $K(\cdot)$ satisfies (i) $\int_{-\infty}^{\infty} K(u)du = 1$, (ii) $\int_{-\infty}^{\infty} u^2 K(u)du < \infty$, and (iii) $K(u) = K(-u)$, and it is twice differentiable with Lipschitz continuous second-order derivative $K''(u)$

[3]See Section 2.5.1. Actually, Serfling (1980) provides a better rate for the remainder.

satisfying (i) $\int_{-\infty}^{\infty} K''(u)udu < \infty$ *and (ii) there exist positive constants* C_1 *and* C_2 *such that* $|K''(u_1) - K''(u_2)| \leq C_2 |u_1 - u_2|^2$ *for* $|u_1 - u_2| \geq C_1$.

We also need the following rate assumption on the bandwidth.

Assumption 2.5. ***Rate Assumption*** *Assume that* $n \uparrow \infty$ *and* $h \downarrow 0$ *such that* $nh^3 \uparrow \infty$.

The non-standard condition $nh^3 \uparrow \infty$ is due to the estimation of y_τ. Since we need to expand $\hat{f}_Y(\hat{y}_\tau) - \hat{f}_Y(y_\tau)$, the derivative of $\hat{f}_Y(y)$ will entail a slower decay for h. The details can be found in the proof of Lemma 2.3. We note, however, that $nh^3 \uparrow \infty$ implies the usual $nh \uparrow \infty$.

The estimator of $f_Y(y)$ is then given by

$$\hat{f}_Y(y) = \frac{1}{n}\sum_{i=1}^{n} K_h(Y_i - y),$$

where $K_h(u) := K(u/h)/h$.

Lemma 2.3. *Let Assumptions 2.4 and 2.5 hold. Then*

$$\hat{f}_Y(y) - f_Y(y) = \mathbb{P}_n \psi_{f_Y}(y) + B_{f_Y}(y) + o_p(h^2),$$

where

$$\psi_{f_Y}(y) := K_h(Y - y) - E[K_h(Y - y)] = O_p(n^{-1/2}h^{-1/2})$$

and

$$B_{f_Y}(y) = \frac{1}{2}h^2 f_Y''^2(y) \int_{-\infty}^{\infty} u^2 K(u)du.$$

Furthermore, for the quantile estimator $\hat{y}_\tau$ *of* y_τ *that satisfies Lemma 2.2, we have*

$$\hat{f}_Y(\hat{y}_\tau) - \hat{f}_Y(y_\tau) = f_Y(\hat{y}_\tau) - f_Y(y_\tau) + R_{f_Y} = f_Y'(y_\tau)\mathbb{P}_n \psi_Q(y_\tau) + R_{f_Y},$$

where

$$R_{f_Y} = o_p(n^{-1/2}h^{-1/2}).$$

In order to isolate the contributions of $\hat{f}$ and $\hat{y}_\tau$, we can use Lemma 2.3 to write

$$\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau) = \hat{f}_Y(y_\tau) - f_Y(y_\tau) + f_Y(\hat{y}_\tau) - f_Y(y_\tau) + R_{f_Y}. \tag{2.22}$$

The first pair of terms on the right-hand side of (2.22) represents the dominant term and reflects the uncertainty in the estimation of f_Y. The second pair of terms reflects the error from estimating y_τ. In order to ensure that $R_{f_Y} = o_p(n^{-1/2}h^{-1/2})$, we need $nh^3 \uparrow \infty$, as stated in Assumption 2.5. We will use (2.22) repeatedly.

2.5.2 Parametric Propensity Score

In this subsection we assume that the propensity score is known up to a finite-dimensional vector α_0. For example, the propensity score is a logit function of $W = (Z, X)$. In this case, we estimate α_0 using the maximum likelihood estimator.

Assumption 2.6. *The propensity score is known up to a finite-dimensional vector $\alpha_0 \in \mathbb{R}^{d_\alpha}$.*

We denote the propensity by $P(Z, X, \alpha_0)$. Under Assumption 2.6, the parameter $M_{\tau,z}$ can be written as

$$M_{\tau,z} = -\frac{1}{f_Y(y_\tau)} \cdot \frac{1}{T_1} \cdot T_2, \tag{2.23}$$

where

$$T_1 = E\left[\frac{\partial P(Z, X, \alpha_0)}{\partial z}\right] \text{ and } T_2 = E\left[\frac{\partial m_0(y_\tau, P(Z, X, \alpha_0), X)}{\partial z}\right].$$

First, we estimate T_1, the average value of the derivative of the propensity score, by

$$T_{1n}(\hat{\alpha}) := \frac{1}{n}\sum_{i=1}^{n} \frac{\partial P(z, x, \hat{\alpha})}{\partial z}\bigg|_{(z,x)=(Z_i,X_i)}.$$

To save space, we slightly abuse notation and write $T_{1n}(\hat{\alpha})$ as

$$T_{1n}(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial P(Z_i, X_i, \hat{\alpha})}{\partial z}. \tag{2.24}$$

We adopt this convention in the rest of the chapter.

Lemma 2.4. *Suppose that*

(i) $\hat{\alpha}$ *admits the linear representation*

$$\hat{\alpha} - \alpha_0 = \mathbb{P}_n \psi_{\alpha_0} + o_p(n^{-1/2}); \tag{2.25}$$

where $\psi_{\alpha_0}(W_i)$ *is a mean-zero* $d_\alpha \times 1$ *random vector with* $E\|\psi_{\alpha_0}(W_i)\|^2 < \infty$*, and* $\|\cdot\|$ *denotes the usual Euclidean norm;*

(ii) the variance of $\frac{\partial P(Z,X,\alpha_0)}{\partial z} := \frac{\partial P(z,x,\alpha_0)}{\partial z}\big|_{(z,x)=(Z,X)}$ *is finite;*

(iii) the $d_\alpha \times 1$ *derivative vector* $\frac{\partial^2 P(Z,X,\alpha)}{\partial \alpha \partial z}$ *exists in an open neighborhood around* α_0;

(iv) the following uniform law of large numbers holds:

$$\sup_{\alpha \in \mathscr{A}_0} \left\| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 P(Z_i, X_i, \alpha)}{\partial \alpha \partial z} - E\left[\frac{\partial^2 P(Z,X,\alpha)}{\partial \alpha \partial z}\right] \right\| \xrightarrow{p} 0$$

where $\mathscr{A}_0$ *is a neighborhood around* α_0*, and*

$$\alpha \mapsto E\left[\frac{\partial^2 P(Z,X,\alpha)}{\partial \alpha \partial z}\right]$$

is continuous on $\mathscr{A}_0$.

Then $T_{1n}(\hat{\alpha})$ *can be represented as*

$$T_{1n}(\hat{\alpha}) - T_1 = E\left[\frac{\partial^2 P(Z,X,\alpha_0)}{\partial z \partial \alpha}\right]' \mathbb{P}_n \psi_{\alpha_0} + \mathbb{P}_n \psi_{\partial P} + o_p(n^{-1/2}),$$

where

$$\psi_{\partial P} := \frac{\partial P(Z,X,\alpha_0)}{\partial z} - E\left[\frac{\partial P(Z,X,\alpha_0)}{\partial z}\right].$$

We can rewrite the main result of Lemma 2.4 as

$$\begin{aligned}T_{1n}(\hat{\alpha})-T_1 &= T_{1n}(\alpha_0)-T_1+T_{1n}(\hat{\alpha})-T_{1n}(\alpha_0)\\ &= T_{1n}(\alpha_0)-T_1+E\left[\frac{\partial P(Z,X,\hat{\alpha})}{\partial z}\right]\\ &-E\left[\frac{\partial P(Z,X,\alpha_0)}{\partial z}\right]+o_p(n^{-1/2}).\end{aligned} \tag{2.26}$$

Equation (2.26) has the same interpretation as equation (2.22). It consists of a pair of leading terms that ignores the estimation uncertainty in $\hat{\alpha}$ but accounts for the variability of the sample mean, and another pair that accounts for the uncertainty in $\hat{\alpha}$ but ignores the variability of the sample mean.

We estimate the second average derivative T_2 by

$$T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha}):=\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\hat{m}(\hat{y}_\tau,P(Z_i,X_i,\hat{\alpha}),X_i)}{\partial z}. \tag{2.27}$$

This can be regarded as a four-step estimator. The first step estimates y_τ, the second step estimates α_0, the third step estimates the conditional expectation $m_0(y,P(Z,X,\alpha_0),X)$ using the generated regressor $P(Z,X,\hat{\alpha})$, and the fourth step averages the derivative (with respect to Z) over X and the generated regressor $P(Z,X,\hat{\alpha})$.

We use the series method to estimate m_0. To alleviate notation, define the vector

$$\tilde{w}(\alpha):=(P(z,x,\alpha),x)' \text{ and } \tilde{W}_i(\alpha):=(P(Z_i,X_i,\alpha),X_i)'.$$

Both $\tilde{w}(\alpha)$ and $\tilde{W}_i(\alpha)$ are in $\mathbb{R}^{d_X+1}$. Let

$$\phi^J(\tilde{w}(\alpha))=(\phi_{1J}(\tilde{w}(\alpha)),\ldots,\phi_{JJ}(\tilde{w}(\alpha)))'$$

be a vector of J basis functions of $\tilde{w}(\alpha)$ with finite second moments. Here each $\phi_{jJ}(\cdot)$ is a differentiable basis function. Then, the conditional expectation estimator is

$$\hat{m}(\hat{y}_\tau, \tilde{w}(\hat{\alpha})) = \phi^J(\tilde{w}(\hat{\alpha}))'\hat{b}(\hat{\alpha}, \hat{y}_\tau), \tag{2.28}$$

where $\hat{b}(\hat{\alpha}, \hat{y}_\tau)$ is the least squares estimate:

$$\hat{b}(\hat{\alpha}, \hat{y}_\tau) = \left(\sum_{i=1}^{n} \phi^J(\tilde{W}_i(\hat{\alpha}))\phi^J(\tilde{W}_i(\hat{\alpha}))' \right)^{-1} \sum_{i=1}^{n} \phi^J(\tilde{W}_i(\hat{\alpha}))\mathbb{1}\left\{Y_i \leq \hat{y}_\tau\right\}. \tag{2.29}$$

The estimator of the derivative is then

$$\frac{\partial \hat{m}(\hat{y}_\tau, \tilde{w}(\hat{\alpha}))}{\partial z} = \frac{\partial \phi^J(\tilde{w}(\hat{\alpha}))'}{\partial z} \hat{b}(\hat{\alpha}, \hat{y}_\tau), \tag{2.30}$$

and the estimator of the average derivative becomes

$$T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \phi^J(\tilde{W}_i(\hat{\alpha}))'}{\partial z} \hat{b}(\hat{\alpha}, \hat{y}_\tau). \tag{2.31}$$

We use the path derivative approach of Newey (1994) to obtain a decomposition of $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2$, which is similar to that in Section 2.1 of Hahn and Ridder (2013). To describe the idea, let

$$O := (Y, Z, X, D)$$

be the vector of observations, and let $\{F_\theta\}$ be a path of distributions indexed by $\theta \in \mathbb{R}$ such that F_{θ_0} is the true distribution of O. The parametric assumption on the propensity score need not be imposed on the path.[4] The score of the parametric submodel is

$$S(O) = \left. \frac{\partial \log dF_\theta(O)}{\partial \theta} \right|_{\theta=\theta_0}.$$

[4]As we show later, the error from estimating the propensity score does not affect the asymptotic variance of $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})$.

For any θ, we define

$$T_{2,\theta} = E_\theta \left[\frac{\partial m_\theta(y_{\tau,\theta}, \tilde{W}(\alpha_\theta))}{\partial z} \right]$$

where m_θ, $y_{\tau,\theta}$, and α_θ are the probability limits of $\hat{m}$, $\hat{y}_\tau$, and $\hat{\alpha}$, respectively, when the distribution of O is F_θ. Note that when $\theta = \theta_0$, we have $T_{2,\theta_0} = T_2$. Suppose the set of scores $\{S(O)\}$ for all parametric submodels $\{F_\theta\}$ can approximate in the mean square any zero-mean, finite-variance function of O.[5] If the function $\theta \to T_{2,\theta}$ is differentiable at θ_0 and we can write

$$\left. \frac{\partial T_{2,\theta}}{\partial \theta} \right|_{\theta=\theta_0} = E\left[\Gamma(O)S(O)\right] \tag{2.32}$$

for some mean-zero and finite second-moment function $\Gamma(\cdot)$ and any path F_θ, then, by Theorem 2.1 of Newey (1994), the asymptotic variance of $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})$ is $E[\Gamma(O)^2]$.

In the next lemma, we will show that $\theta \to T_{2,\theta}$ is differentiable at θ_0. Suppose for the moment this is the case. Then, by the chain rule, we can write

$$\begin{aligned} \left. \frac{\partial T_{2,\theta}}{\partial \theta} \right|_{\theta=\theta_0} &= \left. \frac{\partial}{\partial \theta} E_\theta \left[\frac{\partial m_\theta(y_{\tau,\theta}, \tilde{W}(\alpha_\theta))}{\partial z} \right] \right|_{\theta=\theta_0} \\ &= \left. \frac{\partial}{\partial \theta} E_\theta \left[\frac{\partial m_0(y_\tau, \tilde{W}(\alpha_0))}{\partial z} \right] \right|_{\theta=\theta_0} + \left. \frac{\partial}{\partial \theta} E \left[\frac{\partial m_\theta(y_\tau, \tilde{W}(\alpha_0))}{\partial z} \right] \right|_{\theta=\theta_0} \\ &+ \left. \frac{\partial}{\partial \theta} E \left[\frac{\partial m_0(y_{\tau,\theta}, \tilde{W}(\alpha_0))}{\partial z} \right] \right|_{\theta=\theta_0} + \left. \frac{\partial}{\partial \theta} E \left[\frac{\partial m_0(y_\tau, \tilde{W}(\alpha_\theta))}{\partial z} \right] \right|_{\theta=\theta_0} \end{aligned} \tag{2.33}$$

To use Theorem 2.1 of Newey (1994), we need to write all these terms in an outer-product form, namely the form of the right-hand side of (2.32). To search for the required pathwise derivative of $\Gamma(\cdot)$, we can examine one component of $T_{2,\theta}$ at a time by treating the remaining components as known, an observation due to Newey (1994).

The next lemma provides the conditions under which we can ignore the error from estimating the propensity score in our asymptotic analysis.

Lemma 2.5. *Assume that*

[5] This is the "generality" requirement of the family of distibutions in Newey (1994).

(i) (Z,X) *is absolutely continuous with density* $f_{ZX}(z,x)$ *satisfying*

(a) $f_{ZX}(z,x)$ *is continuously differentiable with respect to* z *in* $\mathscr{Z}\times\mathscr{X}$*;*

(b) for each $x\in\mathscr{X}$*,* $f_{ZX}(z,x)=0$ *for any* z *on the boundary of its support* $\mathscr{Z}(x)$*;*

(c) $\frac{\partial\log f_{Z,X}(Z,X)}{\partial z}$ *has finite second moments;*

(ii) the following conditional mean independence holds:

$$\begin{aligned} E\left[\frac{\partial\log f_W(W)}{\partial z}\middle|\tilde{W}(\alpha_0),\frac{\partial P(W,\alpha_0)}{\partial\alpha_0}\right] &= E\left[\frac{\partial\log f_W(W)}{\partial z}\middle|\tilde{W}(\alpha_0)\right], \\ E\left[\mathbb{1}\{Y\le y_\tau\}|\tilde{W}(\alpha_0),\frac{\partial P(W,\alpha_0)}{\partial\alpha_0}\right] &= E\left[\mathbb{1}\{Y\le y_\tau\}|\tilde{W}(\alpha_0)\right]; \end{aligned}$$

(iii) $m(y_\tau,\tilde{w}(\alpha_0))$ *is continuously differentiable with respect to* z *for all orders, and for a neighborhood* Θ_0 *of* θ_0*, the following holds:*

$$\begin{aligned} &E\left[\sup_{\theta\in\Theta_0}\left|\frac{\partial}{\partial\alpha_\theta}\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_\theta))}{\partial z}\right|\right]<\infty, \\ &E\sup_{\theta\in\Theta_0}\left|\frac{\partial}{\partial\alpha_\theta}E\left[\frac{\partial\log f_W(W)}{\partial z}\middle|\tilde{W}(\alpha_\theta)\right]\right|<\infty, \\ &E\sup_{\theta\in\Theta_0}\left|\frac{\partial}{\partial\alpha_\theta}\left\{m_0(y_\tau,\tilde{W}(\alpha_\theta))E\left[\frac{\partial\log f_W(W)}{\partial z}\middle|\tilde{W}(\alpha_\theta)\right]\right\}\right|<\infty. \end{aligned}$$

Then

$$\frac{\partial}{\partial\theta}E\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_\theta))}{\partial z}\right]\bigg|_{\theta=\theta_0}=0.$$

Condition (ii) of the Lemma holds trivially if the equation system $P(W,\alpha_0)=P(w,\alpha_0)$ and $X=x$ with $W=(Z,X)$ as the unknown has a unique solution $Z=z$, so that the equation system implies $W=w$. In this case, conditioning on $\frac{\partial P(W,\alpha_0)}{\partial\alpha_0}$ becomes redundant. In general, the equation system may not have a unique solution for Z. Condition (ii) will hold if knowing that

$$\frac{\partial P(W,\alpha_0)}{\partial\alpha_0}=\frac{\partial P(w,\alpha_0)}{\partial\alpha_0}$$

does not change the solution set. As an example, we drop X and consider

$$P(Z,\alpha_0) = \frac{\exp\left(Z^2\alpha_0\right)}{1+\exp\left(Z^2\alpha_0\right)}, \text{ so that } \frac{\partial P(Z,\alpha_0)}{\partial\alpha_0} = Z^2\frac{\exp\left(Z^2\alpha_0\right)}{1+\exp\left(Z^2\alpha_0\right)}.$$

For any $z>0$, $P(Z,\alpha_0) = P(z,\alpha_0)$ has two solutions: $Z = z$ and $Z = -z$. The solution set will not change if we also know that

$$\frac{\partial P(Z,\alpha_0)}{\partial\alpha_0} = \frac{\partial P(z,\alpha_0)}{\partial\alpha_0}.$$

Thus Condition (ii) holds.

The next lemma establishes a stochastic approximation of $T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha}) - T_2$ and provides the influence function as well. The assumptions of the Lemma are adapted from Newey (1994). These assumptions are not necessarily the weakest possible.

Lemma 2.6. *Suppose that*

(i) the support $\mathscr{Z}\times\mathscr{X}$ *of* (Z,X) *is* $[z_l,z_u]\times[x_{1l},x_{1u}]\times[x_{2l},x_{2u}]\times\cdots\times[x_{d_Xl},x_{d_Xu}]$, $f_{ZX}(z,x)$ *is bounded below by* $C\times(z-z_l)^\kappa(z_u-z)^\kappa\left(M_{j=1}^{d_X}(x-x_{jl})^\kappa(x_{ju}-x)^\kappa\right)$ *for some* $C>0$ *and* $\kappa>0$, *and*

$$\int_{\mathscr{Z}\times\mathscr{X}}\sup_{\theta\in\Theta_0}\left|\frac{\partial f_{ZX}(z,x;\theta)}{\partial\theta}\right| dzdx < \infty;$$

(ii) there is a constant C such that $\partial^a m(y_\tau,\tilde{w}(\alpha_0))/\partial z^a \leq C^a$ *for all* $a\in\mathbb{N}$;

(iii) the number of series terms, J, *satisfies* $J(n) = O(n^\rho)$ *for some* $\rho>0$, *and* $J^{7+2\kappa} = O(n)$;

(iv) the map $(y,z)\mapsto m_0(y,\tilde{w}(\alpha_0))$ *is differentiable, and*

$$E\left[\sup_{y\in\mathscr{Y}}\left|\frac{m_0(y,\tilde{W}(\alpha_0))}{\partial y\partial z}\right|\right] < \infty;$$

(v) the following stochastic equicontinuity conditions hold:

$$
\begin{aligned}
&E\left[\frac{\partial \hat{m}(y_\tau, \tilde{W}(\alpha_0))}{\partial z} - \frac{\partial m_0(y_\tau, \tilde{W}(\alpha_0))}{\partial z}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial \hat{m}(y_\tau, \tilde{W}_i(\alpha_0))}{\partial z} - \frac{\partial m_0(y_\tau, \tilde{W}_i(\alpha_0))}{\partial z}\right] + o_p(n^{-1/2}), \\
&E\left[\frac{\partial m_0(\hat{y}_\tau, \tilde{W}(\alpha_0))}{\partial z} - \frac{\partial m_0(y_\tau, \tilde{W}(\alpha_0))}{\partial z}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial m_0(\hat{y}_\tau, \tilde{W}_i(\alpha_0))}{\partial z} - \frac{\partial m_0(y_\tau, \tilde{W}_i(\alpha_0))}{\partial z}\right] + o_p(n^{-1/2});
\end{aligned}
$$

(vi) the assumptions of Lemma 2.5 hold.

Then, we have the decomposition

$$
\begin{aligned}
T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 &= T_{2n}(y_\tau, m_0, \alpha_0) - T_2 \\
&+ T_{2n}(y_\tau, \hat{m}, \alpha_0) - T_{2n}(y_\tau, m_0, \alpha_0) \\
&+ T_{2n}(\hat{y}_\tau, m_0, \alpha_0) - T_{2n}(y_\tau, m_0, \alpha_0) \\
&+ o_p(n^{-1/2}). \qquad (2.34)
\end{aligned}
$$

Additionally,

$$
T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 = \mathbb{P}_n \psi_{\partial m_0} - \mathbb{P}_n \psi_{m_0} + \mathbb{P}_n \tilde{\psi}_Q(y_\tau) + o_p(n^{-1/2}),
$$

where

$$
\begin{aligned}
\psi_{\partial m_0} &:= \frac{\partial m_0(y_\tau, \tilde{W}(\alpha_0))}{\partial z} - T_2, \\
\psi_{m_0} &:= \left[\mathbb{1}\{Y \le y_\tau\} - m_0(y_\tau, \tilde{W}(\alpha_0)\right] \times E\left[\frac{\partial \log f_W(W)}{\partial z}\middle| \tilde{W}(\alpha_0)\right],
\end{aligned}
$$

and

$$\begin{aligned}\tilde{\psi}_Q(y_\tau) &:= E\left[\frac{\partial f_{Y|\tilde{W}(\alpha_0)}(y_\tau|\tilde{W}(\alpha_0))}{\partial z}\right]\psi_Q(y_\tau)\\ &= E\left[\frac{\partial f_{Y|\tilde{W}(\alpha_0)}(y_\tau|\tilde{W}(\alpha_0))}{\partial z}\right]\left[\frac{\tau-\mathbb{1}\{Y\le y_\tau\}}{f_Y(y_\tau)}\right].\end{aligned}$$

Lemma 2.6 characterizes the contribution of each stage to the final influence function. The contribution of the estimation of m_0, given by $\mathbb{P}_n\psi_{m_0}$, corresponds to the one in Proposition 5 of Newey (1994).

2.5.3 Estimation of the UQE

We estimate the UQE by

$$\hat{M}_{\tau,z}(\hat{y}_\tau,\hat{f}_Y,\hat{m},\hat{\alpha}) = -\frac{1}{\hat{f}_Y(\hat{y}_\tau)}\frac{T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})}{T_{1n}(\hat{\alpha})}. \tag{2.35}$$

With the asymptotic linear representations of the arguments $\hat{f}_Y(\hat{y}_\tau)$, $T_{1n}(\hat{\alpha})$, and $T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})$, we can obtain the asymptotic linear representation of $\hat{M}_{\tau,z}(\hat{y}_\tau,\hat{f}_Y,\hat{m},\hat{\alpha})$. The next theorem follows from combining Lemmas 2.3, 2.4, 2.5, and 2.6.

Theorem 2.5. *Under the assumptions of Lemmas 2.3, 2.4, 2.5, and 2.6, we have*

$$\begin{aligned}\hat{M}_{\tau,z}-M_{\tau,z} &= \frac{T_2}{f_Y(y_\tau)^2T_1}\left[\mathbb{P}_n\psi_{f_Y}(y_\tau)+B_{f_Y}(y_\tau)\right]+\frac{T_2}{f_Y(y_\tau)^2T_1}f_Y'(y_\tau)\mathbb{P}_n\psi_Q(y_\tau)\\ &+ \frac{T_2}{f_Y(y_\tau)T_1^2}\mathbb{P}_n\psi_{\partial P}+\frac{T_2}{f_Y(y_\tau)T_1^2}E\left[\frac{\partial^2P(Z,X,\alpha_0)}{\partial z\partial\alpha_0'}\right]\mathbb{P}_n\psi_{\alpha_0}\\ &- \frac{1}{f_Y(y_\tau)T_1}\mathbb{P}_n\psi_{\partial m_0}+\frac{1}{f_Y(y_\tau)T_1}\mathbb{P}_n\psi_{m_0}\\ &- \frac{1}{f_Y(y_\tau)T_1}\mathbb{P}_n\tilde{\psi}_Q(y_\tau)+R_M,\end{aligned} \tag{2.36}$$

where

$$\begin{aligned} R_M &= O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|^2\right)+O_p\left(n^{-1}\right)+O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right) \\ &+ O_p\left(|R_{f_Y}|\right)+o_p(n^{-1/2})+o_p(h^2). \end{aligned} \tag{2.37}$$

Furthermore, under Assumption 2.5, $\sqrt{nh}R_M = o_p(1)$.

Equation (2.36) consists of six influence functions and a bias term. The bias term $B_{f_Y}(y_\tau)$ arises from estimating the density and is of order $O(h^2)$. The six influence functions reflect the impact of each estimation stage. The rate of converge of $\hat{M}_{\tau,z}$ is slowed down through $\mathbb{P}_n\psi_{f_Y}(y_\tau)$, which is of order $O_p(n^{-1/2}h^{-1/2})$. We can summarize the results of Theorem 2.5 in a single equation:

$$\hat{M}_{\tau,z}-M_{\tau,z}=\mathbb{P}_n\psi_{M_\tau}+\tilde{B}_{f_Y}(y_\tau)+o_p(n^{-1/2}h^{-1/2}),$$

where ψ_{M_τ} collects all the influence functions in (2.36) except for the bias, and

$$\tilde{B}_{f_Y}(y_\tau):=\frac{T_2}{f_Y(y_\tau)^2T_1}B_{f_Y}(y_\tau).$$

The bias term is $o(n^{-1/2}h^{-1/2})$ by Assumption 2.5. The following corollary provides the asymptotic distribution of $\hat{M}_{\tau,z}$.

Corollary 2.6. *Under the assumptions of Theorem 2.5,*

$$\sqrt{nh}\left(\hat{M}_{\tau,z}-M_{\tau,z}\right)=\sqrt{n}\mathbb{P}_n\sqrt{h}\psi_{M_\tau}+o_p(1)\Rightarrow\mathcal{N}(0,V_\tau),$$

where

$$V_\tau=\lim_{h\downarrow 0}E\left[h\psi_{M_\tau}^2\right]. \tag{2.38}$$

From the perspective of asymptotic theory, all the terms $\sqrt{nh}\mathbb{P}_n\psi_Q(y_\tau)$, $\sqrt{nh}\mathbb{P}_n\psi_{\partial P}$, $\sqrt{nh}\mathbb{P}_n\psi_{\alpha_0}$, $\sqrt{nh}\mathbb{P}_n\psi_{\partial m_0}$, and $\sqrt{nh}\mathbb{P}_n\psi_{m_0}$ are of order $O_p(h)=o_p(1)$ and hence can be ignored in large samples. The asymptotic variance is then given by

$$V_\tau = \frac{T_2}{f_Y(y_\tau)^2 T_1}\lim_{h\downarrow 0} h\psi_f^2(y_\tau,h) = \frac{T_2}{f_Y(y_\tau)T_1}\int_{-\infty}^{\infty} K^2(u)\,du.$$

However, V_τ ignores all estimation uncertainties except that in $\hat{f}_Y(y_\tau)$, and we do not expect it to reflect the finite-sample variability of $\sqrt{nh}(\hat{M}_{\tau,z} - M_{\tau,z})$ well. To improve the finite-sample performances, we keep the dominating term from each source of estimation errors and employ a sample counterpart of $Eh\psi_{M_\tau}^2$ to estimate V_τ. This is done in the next subsection.

2.5.4 Estimation of the Asymptotic Variance of the UQE Estimator

The asymptotic variance V_τ in (2.38) can be estimated by the plug-in estimator

$$\hat{V}_\tau = \frac{h}{n}\sum_{i=1}^{n}\hat{\psi}_{M_\tau,i}^2, \tag{2.39}$$

where, by Theorem 2.5,

$$\begin{aligned}
\hat{\psi}_{M_\tau,i} &= \frac{\hat{T}_{2n}}{\hat{f}_Y(\hat{y}_\tau)^2\hat{T}_{1n}}\hat{\psi}_{f,i}(\hat{y}_\tau) + \frac{\hat{T}_{2n}}{\hat{f}_Y(y_\tau)^2\hat{T}_{1n}}\hat{f}_Y'(y_\tau)\hat{\psi}_{Q,i}(\hat{y}_\tau) \\
&+ \frac{\hat{T}_{2n}}{\hat{f}_Y(\hat{y}_\tau)\hat{T}_{1n}^2}\hat{\psi}_{\partial P,i} + \frac{\hat{T}_{2n}}{\hat{f}_Y(\hat{y}_\tau)\hat{T}_{1n}^2}\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial^2 P(W_i,\hat{\alpha})}{\partial z\partial\hat{\alpha}'}\right]\hat{\psi}_{\alpha,i} \\
&- \frac{1}{\hat{f}_Y(\hat{y}_\tau)\hat{T}_{1n}}\hat{\psi}_{\partial m,i} + \frac{1}{\hat{f}_Y(\hat{y}_\tau)\hat{T}_{1n}}\hat{\psi}_{m,i} \\
&- \frac{1}{\hat{f}_Y(\hat{y}_\tau)\hat{T}_{1n}}\hat{E}\left[\frac{\partial f_{Y|\tilde{W}(\alpha_0)}(y_\tau|\tilde{W}(\hat{\alpha}))}{\partial z}\right]\hat{\psi}_{Q,i}(\hat{y}_\tau).
\end{aligned}$$

In this equation, $\hat{T}_{2n} = T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})$, $\hat{T}_{1n} = T_{1n}(\hat{\alpha})$,

$$\begin{aligned}
\hat{\psi}_{f,i}(\hat{y}_\tau) &= K_h\left([Y_i - \hat{y}_\tau]\right) - \frac{1}{n}\sum_{j=1}^{n} K_h\left(\left[Y_j - \hat{y}_\tau\right]\right) \\
\hat{\psi}_{Q,i}(\hat{y}_\tau) &= \frac{\tau - \mathbb{1}\{Y_i \le \hat{y}_\tau\}}{\hat{f}_Y(\hat{y}_\tau)}, \\
\hat{\psi}_{\partial P,i} &= \frac{\partial P(W_i,\hat{\alpha})}{\partial z} - \frac{1}{n}\sum_{j=1}^{n} \frac{\partial P(W_j,\hat{\alpha})}{\partial z}, \\
\hat{\psi}_{\alpha,i} &= \left(-\frac{1}{n}\sum_{i=1}^{n} \frac{[P_\partial(W_i,\hat{\alpha})]^2 W_i' W_i}{P(W_i,\hat{\alpha})[1-P(W_i,\hat{\alpha})]}\right)^{-1} \frac{P_\partial(W_i,\hat{\alpha}) W_i [D_i - P(W_i,\hat{\alpha})]}{P(W_i,\hat{\alpha})[1-P(W_i,\hat{\alpha})]}, \\
\hat{\psi}_{\partial m,i} &= \frac{\partial \hat{m}(y_\tau, \tilde{W}_i(\hat{\alpha}))}{\partial z} - \hat{T}_{2n} \\
\hat{\psi}_{m,i} &= \left(\mathbb{1}\{Y_i \le \hat{y}_\tau\} - m_0(\hat{y}_\tau, \tilde{W}_i(\hat{\alpha}))\right) \times \hat{E}\left[\left.\frac{\partial \log f_W(W_i)}{\partial z}\right| \tilde{W}_i(\hat{\alpha})\right],
\end{aligned}$$

and

$$\hat{E}\left[\frac{\partial f_{Y|\tilde{W}(\alpha_0)}(y_\tau|\tilde{W}(\hat{\alpha}))}{\partial z}\right] = \frac{1}{n}\sum_{j=1}^{n} \frac{\partial}{\partial z} \left.\frac{\sum_{i\neq j} K_h([Y_i - y_\tau]) \cdot \mathbb{K}_h\left[\tilde{W}_i(\hat{\alpha}) - \tilde{w}(\hat{\alpha})\right]}{\sum_{i\neq j} \mathbb{K}_h\left[\tilde{W}_i(\hat{\alpha}) - \tilde{w}(\hat{\alpha})\right]}\right|_{w=W_j}$$

for the rescaled kernel function $\mathbb{K}_h(\cdot)$ defined by

$$\mathbb{K}_h([\chi_1,\ldots,\chi_\ell]) = \prod_{j=1}^{\ell} K_h(\chi_\ell) \text{ with } [\chi_1,\ldots,\chi_\ell] \in \mathbb{R}^\ell.$$

Most of these plug-in estimates are self-explanatory. For example, $\hat{\psi}_{\alpha,i}$ is the estimated influence function for the MLE when $P(W_i,\hat{\alpha}) = P(W_i\hat{\alpha})$ and $P_\partial(a) = \partial P(a)/\partial a$. If the propensity score function does not take a linear index form, then we need to make some adjustment to $\hat{\psi}_{\alpha,i}$. The only thing we need to do is find the influence function for the MLE, which is an easy task, and then plug $\hat{\alpha}$ into the influence function.

The only remaining quantity that needs some explanation is $\hat{\psi}_{m,i}$, which involves a nonparametric regression of $\frac{\partial \log f_W(W_i)}{\partial z}$ on $\tilde{W}_i(\hat{\alpha}) := [P(Z_i, X_i, \hat{\alpha}), X_i]$. We let

$$\hat{E}\left[\frac{\partial \log f_W(W_i)}{\partial z}\bigg|\tilde{W}_i(\hat{\alpha})\right] = -\phi^J(\tilde{W}_i(\hat{\alpha}))\left(\sum_{\ell=1}^{n}\phi^J(\tilde{W}_\ell(\hat{\alpha}))\phi^J(\tilde{W}_\ell(\hat{\alpha}))'\right)^{-1} \times \sum_{\ell=1}^{n}\frac{\partial \phi^J(\tilde{W}_\ell(\hat{\alpha}))}{\partial z}.$$

To see why this may be consistent for $E\left[\frac{\partial \log f(W_i)}{\partial z}\Big|\tilde{W}_i(\alpha_0)\right]$, we note that under some conditions the following hold:

$$\left\|\frac{1}{n}\sum_{\ell=1}^{n}\frac{\partial \phi^J(\tilde{W}_\ell(\alpha_0))}{\partial z} - E\frac{\partial \phi^J(\tilde{W}(\alpha_0))}{\partial z}\right\|_2 = o_p(1),$$
$$\left\|\frac{1}{n}\sum_{\ell=1}^{n}\phi^J(\tilde{W}_\ell(\alpha_0))\frac{\partial \log f_W(W_\ell)}{\partial z} - E\left[\phi^J(\tilde{W}(\alpha_0))\frac{\partial \log f_W(W)}{\partial z}\right]\right\|_2 = o_p(1).$$

But

$$\begin{aligned} E\frac{\partial \phi^J(\tilde{W}(\alpha_0))}{\partial z} &= \int_{\mathscr{W}}\frac{\partial \phi^J(\tilde{w}(\alpha_0))}{\partial z}f_W(w)\,dw = -\int_{\mathscr{Z}\times\mathscr{X}}\phi^J(\tilde{w}(\alpha_0))\frac{\partial f_W(w)}{\partial z}dzdx \\ &= -\int_{\mathscr{Z}\times\mathscr{X}}\phi^J(\tilde{w}(\alpha_0))\frac{\partial \log f_W(w)}{\partial z}f_W(w)\,dw \\ &= -E\left[\phi^J(\tilde{W}(\alpha_0))\frac{\partial \log f_W(W)}{\partial z}\right]. \end{aligned}$$

Hence

$$\left\|\frac{1}{n}\sum_{\ell=1}^{n}\frac{\partial \phi^J(\tilde{W}_\ell(\alpha_0))}{\partial z} + \frac{1}{n}\sum_{\ell=1}^{n}\phi^J(\tilde{W}_\ell(\alpha_0))\frac{\partial \log f_W(W_\ell)}{\partial z}\right\|_2 = o_p(1),$$

and $\hat{E}\left[\frac{\partial \log f_W(W_i)}{\partial z}\Big|\tilde{W}_i(\hat{\alpha})\right]$ is approximately equal to

$$\phi^J(\tilde{W}_i(\hat{\alpha}))\left(\sum_{\ell=1}^{n}\phi^J(\tilde{W}_\ell(\hat{\alpha}))\phi^J(\tilde{W}_\ell(\hat{\alpha}))'\right)^{-1}\frac{1}{n}\sum_{\ell=1}^{n}\phi^J(\tilde{W}_\ell(\hat{\alpha}))\frac{\partial \log f_W(W_\ell)}{\partial z},$$

which is just a series approximation to $E\left[\frac{\partial \log f_W(W_i)}{\partial z}\Big|\tilde{W}_i(\alpha_0)\right]$.

The consistency of $\hat{V}_\tau$ can be established by using the uniform law of large numbers. The arguments are standard but tedious. We omit the details here.

2.5.5 Testing the Null of No Effect

We can use Corollary 2.6 for hypothesis testing on $M_{\tau,z}$. Since $\hat{M}_{\tau,z}$ converges to $M_{\tau,z}$ at a nonparametric rate, in general the test will have power only against a departure of a nonparametric rate. However, if we are interested in testing the null of a zero effect, that is, $H_0 : M_{\tau,z} = 0$ vs. $H_1 : M_{\tau,z} \neq 0$, we can detect a parametric rate of departure from the null. The reason is that, by (2.23), $M_{\tau,z} = 0$ if and only if $T_2 = 0$, and T_2 can be estimated at the usual parametric rate. Hence, instead of testing $H_0 : M_{\tau,z} = 0$ vs. $H_1 : M_{\tau,z} \neq 0$, we test the equivalent hypotheses $H_0 : T_2 = 0$ vs. $H_1 : T_2 \neq 0$.

Our test is based on the estimator $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})$ of T_2. In view of its influence function given in Lemma 2.6, we can estimate the asymptotic variance of $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})$ by

$$\hat{V}_2 = \frac{1}{n} \sum_{i=1}^{n} \hat{\psi}_{2i}^2,$$

where

$$\begin{aligned}
\hat{\psi}_{2i} \quad := \quad & \frac{\partial \phi^J(\tilde{W}_i(\hat{\alpha}))'}{\partial z} \hat{b}(\hat{\alpha}, \hat{y}_\tau) - T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) \\
- \quad & \left[\sum_{\ell=1}^{n} \frac{\partial \phi^J(\tilde{W}_\ell(\hat{\alpha}))'}{\partial z} \right] \left[\sum_{\ell=1}^{n} \phi^J(\tilde{W}_\ell(\hat{\alpha})) \phi^J(\tilde{W}_\ell(\hat{\alpha}))' \right]^{-1} \\
\times \quad & \phi^J(\tilde{W}_i(\hat{\alpha}))' \left(\mathbb{1}\{Y_i \leq \hat{y}_\tau\} - \phi^J(\tilde{W}_i(\hat{\alpha}))' \hat{b}(\hat{\alpha}, \hat{y}_\tau) \right) \\
+ \quad & \hat{E}\left[\frac{\partial \log f_W(W_i)}{\partial z} \middle| \tilde{W}_i(\hat{\alpha}) \right] \cdot \frac{\tau - \mathbb{1}\{Y_i \leq \hat{y}_\tau\}}{\hat{f}_Y(\hat{y}_\tau)}.
\end{aligned}$$

We can then form the test statistic:

$$\mathcal{T}_2 := \sqrt{n} \frac{T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})}{\sqrt{\hat{V}_2}}.$$

By Lemma 2.6 and using standard arguments, we can show that $T_2 \Rightarrow \mathcal{N}(0,1)$. To save space, we omit the details here.

2.6 Estimation and Inference under a Nonparametric Propensity Score

In this section, we drop Assumption 2.6 and, using the series method, estimate the propensity score non-parametrically. With respect to the results of the previous sections, we only need to modify Lemma 2.4, since Lemma 2.6 shows that we do not need to account for the error from estimating the propensity score.

Let $\hat{P}(w)$ denote the nonparametric series estimator of $P(w)$. The estimator of $T_1 := E\left[\frac{\partial P(W)}{\partial z}\right]$ is now

$$T_{1n}(\hat{P}) := \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \hat{P}(w)}{\partial z}\bigg|_{w=W_i}.$$

The estimator of T_2 is the same as in (2.27) but with $P(W_i, \hat{\alpha})$ replaced by $\hat{P}(W_i)$:

$$T_{2n}(\hat{y}_\tau, \hat{m}, \hat{P}) := \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \hat{m}(\hat{y}_\tau, \hat{P}(W_i), X_i)}{\partial z}, \tag{2.40}$$

where, as in (2.27), $\hat{m}$ is the series estimator of m. The formula is the same as before, and we only need to replace $P(W_i, \hat{\alpha})$ by $\hat{P}(W_i)$. The UQE estimator becomes

$$\begin{aligned} \hat{M}_{\tau,z}(\hat{y}_\tau, \hat{f}_Y, \hat{m}, \hat{P}) &:= -\frac{1}{\hat{f}_Y(\hat{y}_\tau)}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \hat{P}(W_i)}{\partial z}\right]^{-1}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \hat{m}(\hat{y}_\tau, \hat{P}(W_i), X_i)}{\partial z} \\ &= -\frac{1}{\hat{f}_Y(\hat{y}_\tau)}\frac{T_{2n}(\hat{y}_\tau, \hat{m}, \hat{P})}{T_{1n}(\hat{P})}. \end{aligned} \tag{2.41}$$

The following lemma follows directly from Theorem 7.2 of Newey (1994).

Lemma 2.7. *Let Assumption (i) of Lemmas 2.5 and Assumptions (i) and (iii) of Lemma 2.6 hold. Assume further that $P(z,x)$ is continuously differentiable with respect to z for all orders, and that* *there is a constant C such that* $\partial^a P(z,x)/\partial z^a \leq C^a$ *for all* $a \in \mathbb{N}$*. Then*

$$T_{1n}(\hat{P}) - T_1 = \mathbb{P}_n \psi_{\partial P_s} - \mathbb{P}_n \psi_{P_s} + o_p(n^{-1/2}),$$

where we define

$$\psi_{\partial P_s} := \frac{\partial P(W)}{\partial z} - T_1$$

and

$$\psi_{P_s} := (D - P(W)) \times \frac{\partial \log f_W(W)}{\partial z}.$$

Using a proof similar to that of Lemma 2.6, we can show that the influence functions for $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{P})$ and $T_{2n}(\hat{y}_\tau, \hat{m}, P)$ are the same. That is, we have

$$T_{2n}(\hat{y}_\tau, \hat{m}, \hat{P}) - T_2 = \mathbb{P}_n \psi_{\partial m_0} - \mathbb{P}_n \psi_{m_0} + \mathbb{P}_n \tilde{\psi}_Q(y_\tau) + o_p(n^{-1/2}),$$

where

$$\begin{aligned} \psi_{\partial m_0} &:= \frac{\partial m_0(y_\tau, P(W), X)}{\partial z} - T_2, \\ \psi_{m_0} &:= (\mathbb{1}\{Y \leq y_\tau\} - m_0(y_\tau, P(W), X)) \times E\left[\frac{\partial \log f_W(W)}{\partial z} \middle| P(W), X\right], \end{aligned}$$

and

$$\tilde{\psi}_Q(y_\tau) = E\left[\frac{\partial f_{Y|P(W),X}(y_\tau | P(W), X)}{\partial z}\right] \psi_Q(y_\tau).$$

Given the asymptotic linear representations of $T_{1n}(\hat{P}) - T_1$ and $T_{2n}(\hat{y}_\tau, \hat{m}, \hat{P}) - T_2$, we can directly use Lemma 2.6, together with Lemma 2.3, to obtain an asymptotic linear representation of $\hat{M}_{\tau,z}(\hat{y}_\tau, \hat{f}_Y, \hat{m}, \hat{P})$.

Theorem 2.7. *Under the assumptions of Lemmas 2.3, 2.6, and 2.7, we have*

$$
\begin{aligned}
\hat{M}_{\tau,z} - M_{\tau,z} &= \frac{T_2}{f_Y(y_\tau)^2 T_1}\left[\mathbb{P}_n \psi_{f_Y}(y_\tau) + B_{f_Y}(y_\tau)\right] + \frac{T_2}{f_Y(y_\tau)^2 T_1} f_Y'(y_\tau)\mathbb{P}_n \psi_Q(y_\tau) \\
&+ \frac{T_2}{f_Y(y_\tau) T_1^2}\mathbb{P}_n \psi_{\partial P_s} - \frac{T_2}{f_Y(y_\tau) T_1^2}\mathbb{P}_n \psi_{P_s} \\
&- \frac{1}{f_Y(y_\tau) T_1}\mathbb{P}_n \psi_{\partial m_0} + \frac{1}{f_Y(y_\tau) T_1}\mathbb{P}_n \psi_{m_0} \\
&- \frac{1}{f_Y(y_\tau) T_1}\mathbb{P}_n \tilde{\psi}_Q(y_\tau) + R_M, \qquad (2.42)
\end{aligned}
$$

where

$$
\begin{aligned}
R_M &= O_p\left(|\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)|^2\right) + O_p\left(n^{-1}\right) + O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)|\right) \\
&+ O_p\left(|R_{f_Y}|\right) + o_p(n^{-1/2}) + o_p(h^2). \qquad (2.43)
\end{aligned}
$$

Furthermore, under Assumption 2.5, $\sqrt{nh}R_M = o_p(1)$.

We summarize the results of Theorem 2.7 in a single equation:

$$\hat{M}_{\tau,z} - M_{\tau,z} = \mathbb{P}_n \psi_{M_\tau} + \tilde{B}_{f_Y}(y_\tau) + o_p(n^{-1/2}h^{-1/2}),$$

where ψ_{M_τ} collects all the influence functions in (2.42) except for the bias, R_M is absorbed in the $o_p(n^{-1/2}h^{-1/2})$ term, and

$$\tilde{B}_{f_Y}(y_\tau) := \frac{T_2}{f_Y(y_\tau)^2 T_1} B_{f_Y}(y_\tau).$$

The bias term is $o_p(n^{-1/2}h^{-1/2})$ by Assumption 2.5. The following corollary provides the asymptotic distribution of $\hat{M}_{\tau,z}$.

Corollary 2.8. *Under the assumptions of Theorem 2.7,*

$$\sqrt{nh}\left(\hat{M}_{\tau,z} - M_{\tau,z}\right) = \sqrt{n}\mathbb{P}_n \sqrt{h}\psi_{M_\tau} + o_p(1) \Rightarrow \mathcal{N}(0, V_\tau),$$

where

$$V_\tau = \lim_{h \downarrow 0} E\left[h\psi^2_{M_\tau}\right].$$

The asymptotic variance takes the same form as the asymptotic variance in Corollary 2.6. Estimating the asymptotic variance and testing for a zero unconditional effect are entirely similar to the case with a parametric propensity score. We omit the details to avoid repetition and redundancy. From the perspective of implementation, there is no substantive difference between a parametric approach and a nonparametric approach to the propensity score estimation.

2.7 Empirical Application

We estimate the unconditional quantile effect of expanding college enrollment on (log) wages. The outcome variable Y is the log wage, and the binary treatment is the college enrollment status. Thus $p = \Pr(D = 1)$ is the proportion of individuals who ever enrolled in a college. Arguably, the cost of tuition (Z) is an important factor that affects the college enrollment status but not the wage. In order to alter the proportion of enrolled individuals, we consider a policy that subsidizes tuition by a certain amount. The UQE is the effect of this policy on the different quantiles of the unconditional distribution of wages when the subsidy goes to zero. This is the effect that we denote UQE under instrument intervention in Section 2.2. This policy shifts Z, the tuition, to $Z_\delta = Z + s_z(\delta)$ for some $s_z(\delta)$, which is the same for all individuals, and induces the college enrollment to increase from p to $p+\delta$. Note that we do not need to specify $s_z(\delta)$ because we look at the limiting version as $\delta \to 0$.

We use the same data as in Carneiro et al. (2010) and Carneiro et al. (2011): a sample of white males from the 1979 National Longitudinal Survey of Youth (NLSY1979). The web appendix to Carneiro et al. (2011) contains a detailed description of the variables. The outcome variable Y is the log wage in 1991. The treatment indicator D is equal to 1 if the individual ever enrolled in college by 1991, and 0 otherwise. The other covariates are AFQT score, mother's education, number of siblings, average log earnings 1979–2000 in the county of residence at age 17, average unemployment 1979–2000 in the state of residence at age 17, urban residence dummy at age 14, cohort dummies, years of experience in 1991, average local log earnings in 1991, and local unemployment in 1991. We collect these variables into a vector and denote it by X^O.

We assume that the following four variables (denoted by Z_1, Z_2, Z_3, Z_4) enter the selection equation but not the outcome equation: presence of a four-year college in the county of residence at age 14, local earnings at age 17, local unemployment at age 17, and tuition at local public four-year colleges at age 17. The total sample size is 1747, of which 882 individuals had never enrolled in a college ($D = 0$) by 1991, and 865 individuals had enrolled in a college by 1991 ($D = 1$). We compute the UQE of a marginal shift in the tuition at local public four-year colleges at age 17. So in our notation, $Z = Z_4$, $X = (Z_1, Z_2, Z_3, X^O)$, and $W = (Z, X)$.

To estimate the propensity score, we use a parametric logistic specification. To estimate the conditional expectation, we use a series regression using both the estimated propensity score and the covariates X as the regressors. Due to the large number of variables involved, a penalization of $\lambda = 10^{-4}$ was imposed on the L_2-norm of the coefficients, excluding the constant term as in ridge regressions. We compute the UQE at the quantile level $\tau = 0.1, 0.15, \ldots, 0.9$. For each τ, we also construct the 95% (pointwise) confidence interval.

Figure 2.1 presents the results. The UQE ranges between 0.22 and 0.47 across the quantiles, and averages to 0.37. When we estimate the unconditional mean effect, we obtain an estimate of 0.21, which is somewhat consistent with the quantile cases. We interpret these estimates in the following way: a marginal additive change in tuition produces a marginal change in college enrollment which in turn produces a marginal change in (log) wages between 0.22 and 0.47 across quantiles.

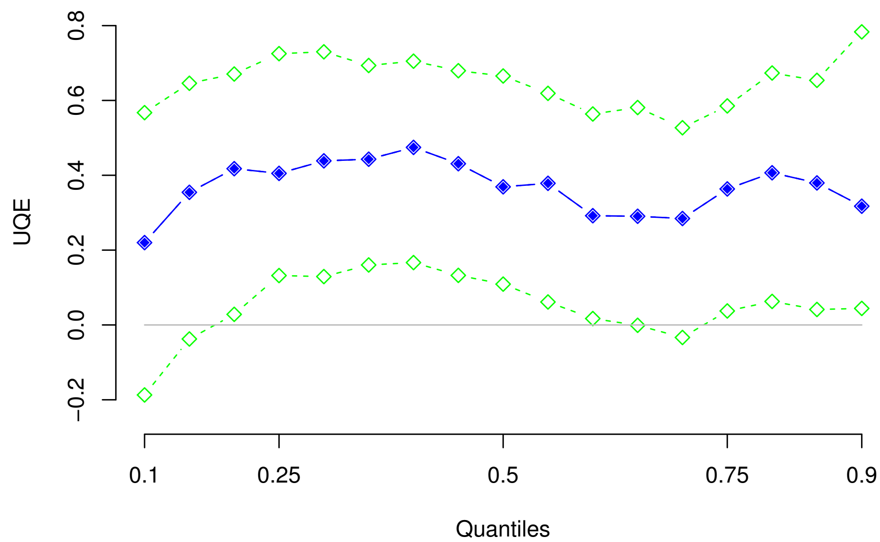

Figure 2.1. Unconditional Quantile Effects.

2.8 Conclusion

In this chapter we study the unconditional policy effect with an endogenous binary treatment. For concreteness, we focus on the unconditional quantile effect, but the basic ideas and insights are applicable if the policy goals are other features of the unconditional distribution of an outcome variable.

We show that when an instrumental variable is available, it is possible to recover the unconditional effect through an application of the local instrumental variable technique. Framing the selection equation as a threshold-crossing model allows us to introduce a new class of marginal treatment effects and represent the unconditional effect as a weighted average of these marginal treatment effects. We find that the unconditional quantile effect and the marginal policy relevant treatment effect can be seen as part of the same family of effects. It is possible to view the latter as a robust version of the former. Both of them are examples of a general unconditional policy effect. To the best of our knowledge, this connection has not been established in either literature.

Chapter 2, in part, is being prepared for submission for publication, and is coauthored with Yixiao Sun. The dissertation author, Julián Martínez-Iriarte was the primary author of this chapter. This research was conducted with restricted access to Bureau of Labor Statistics (BLS) data. The views expressed here do not necessarily reflect the views of the BLS.

Chapter 3

Sensitivity Analysis in Unconditional Quantile Effects

3.1 Introduction

In this chapter we propose a sensitivity analysis on the effect of counterfactual policies that change the proportion of treated individuals. Consider a situation where a policy maker is interested in treating non-treated individuals. The key identification challenge is that we do not the observe the counterfactual outcome of individuals who switch groups, that is, the *newly* treated individuals. In some cases, however, it is still possible to recover the distribution of the unobserved counterfactual outcome. For example, suppose that treatment status is randomly assigned, and a policy maker increases the proportion of treated individuals by randomly selecting non-treated individuals.[1] Although we do not observe the counterfactual outcome of the *newly* treated individuals, we know it is drawn from the same distribution as the *already* treated individuals. Hence, we identified the counterfactual distribution of *newly* treated individuals.

When treatment status is *not* randomly assigned in the first place, the identification strategy previously described breaks down. The reason is that due to the selection bias in the original treatment status, a random selection of individuals from the control group will be drawn from a different distribution. Thus, in the presence of selection bias, identification of the counterfactual distribution requires that the policy maker has enough information to device a

[1] We assume full compliance in both randomizations.

policy such that the (unobservable) distribution of the *newly* treated "matches" the distribution of the *already* treated individuals. This is usually unfeasible. Even if the policy maker has this information, such as when treatment status is randomly assigned, they might not be interested in a policy that matches the previous selection patterns.

The previous discussion highlights that identification of counterfactual distributions results in either very stringent information requirements, or in policies that might not be interesting for the policy maker. In both cases, the distribution of the *newly* treated individuals is restricted. From the point of view of the policy maker, this can rule out many interesting policies. To see this, consider the following example. A policy maker might like to know if an increase in the unionization rate reduces inequality. If unionized workers are relatively high-skilled, and a policy expands unionization with low-skilled workers, then the distribution of wages conditional on being in the union is likely to change.

In order to analyze a richer set of counterfactual policies we impose no restrictions on the distribution of the *newly* treated individuals, and provide partial identification results for two effects: the first one is a global effect that compares the quantiles of the observed outcome, to those of the counterfactual outcome, where the proportion of treated individuals has been increased by δ; and the second one is a marginal effect where we let δ go to zero, and analyze its limiting effect on the unconditional quantiles of the outcome.

The second important contribution of this chapter is to propose a framework for a sensitivity analysis on certain conclusions of interest. To do this, we quantify the departure from point identification by the vertical distance between the distributions of the *newly* treated individuals and the *already* treated individuals. We introduce a curve called the *quantile breakdown frontier*, which quantifies the maximum departure from point identification such that a given set of conclusions holds across different quantiles. Next, we bound the global effects curve using this maximum departure derived from the quantile breakdown frontier. In this way, we obtain an identified region for the global effect curve consistent with the desired conclusions.

The departure from point identification is due to the selection bias induced by the counterfactual policy being different from the original selection bias. We call this difference the *policy selection bias*. The usual selection bias states that treated and non-treated individuals are different in a sense, and that is what explains the selection in the first place. Instead, the policy selection bias is the difference between the distributions of the *newly* treated individuals and the *already* treated individuals. Returning to the unionization example, the policy selection bias arises because the union wages of *newly* unionized workers may not be drawn from the same distribution of the *already* unionized workers. Since we do not know the distribution of union wages of newly unionized workers, hence we can only partially identify the global and marginal effects.

The policy selection bias can be non-negligible even if the original selection into treatment is randomly assigned. The reason is that, for the policy selection bias, what matters is who the *newly* treated individuals are. Conversely, if there is selection bias initially, but the distribution of the *newly* treated "matches" the distribution of the *already* treated individuals, then there will be no policy selection bias. Thus, the policy selection bias depends on the particular counterfactual policy being analyzed, not whether there is selection bias in the original selection mechanism.

It is important to highlight that we do not estimate a quantile treatment effect. The quantile treatment effect is the difference between the τ-quantile under treatment and the τ-quantile under control, and depends on the distribution of the covariates. In a recent contribution, Hsu et al. (2020) investigate the changes in this effect when the distribution of the covariates is manipulated. Aside from treatment status, we do not manipulate the distribution of covariates.

Estimation of both the quantile breakdown frontier and the bounds on the global effect are based on empirical distribution functions and empirical quantiles, and are $\sqrt{n}$-consistent. Inference is more challenging, though. The reason is that the bounds derived from the quantile breakdown frontier are not a fully Hadamard differentiable function of the underlying distributions; there are a few kinks where differentiability fails. However, directional differentiability holds, and we can still exploit the functional Delta method to obtain asymptotic distributions. Since these limiting laws are not Gaussian the standard or "naive" bootstrap is not valid, as shown in Fang and Santos (2019). Instead, we resort to the numerical bootstrap/Delta method of Hong and Li (2018) to construct pointwise confidence intervals and uniform confidence bands.

We apply these methods to the study of unions and inequality, which has long been of interest to labor economists. A recent comprehensive review of this extensive literature is provided by Farber et al. (2020). Using the data in Firpo et al. (2009), our empirical application considers the effect of expanding unionization on the quantiles of the distribution of (log) wages. Using the tools developed in this chapter, we can quantify the amount of policy selection bias that is consistent with a policy that increases the unionization rate by unionizing low earnings workers. By looking at the global effect in the 20^{th} and 80^{th} quantiles of the distribution of wages we investigate the amount of policy selection bias consistent with unions reducing overall inequality. To this end, we ask two questions: whether the 20^{th} quantile increases by more than 10%, and whether the 80^{th} quantile increases less than 10%. We find that these changes are consistent with moderate values of policy selection bias.

The chapter is organized as follows: Section 3.2 reviews the literature; Section 3.3 introduces our framework and shows how to construct the identified regions; Section 3.4 introduces the quantile breakdown frontier and explains the sensitivity analysis procedure; Section 3.5 discusses estimation and inference; Section 3.6 contains the empirical application; and Section 3.7 concludes. We relegate all proofs to Appendix.

3.2 Related Literature

There is an extensive literature devoted to the analysis of counterfactual distributions. A good reference is Firpo et al. (2011). In this chapter, we focus on counterfactual distributions that arise as a result of a counterfactual policy that changes the proportion of treated individuals. The Policy Relevant Treatment Effect (PRTE) of Heckman and Vytlacil (2001b, 2005), and the Marginal PRTE (MPRTE) of Carneiro et al. (2010, 2011) are examples of the aforementioned global and marginal effects. The difference is that they analyze the unconditional mean of the outcome. Identification relies on the a separable threshold model for the selection equation, and the availability of a continuous instrumental variable. In their setting, the proportion of treated individuals is changed by manipulating the instrumental variable. Our analysis does not make any assumptions on the selection equation. We do not require an instrumental variable either.

Our sensitivity analysis is based on the breakdown analysis of Kline and Santos (2013) and Masten and Poirier (2020). Kline and Santos (2013) perform a sensitivity analysis in a different context: departures from a missing (data) at random assumption. In a manner similar to us, this departure is measured as the Kolmogorov-Smirnov distance between the distribution of observed outcomes and the (unobserved) distribution of missing outcomes. Our quantile breakdown frontier builds on the breakdown frontier introduced by Masten and Poirier (2020). However, the quantile breakdown frontier takes advantage of the unique features of the policy selection bias: for each quantile the breakdown frontier of Masten and Poirier (2020) is a rectangle. This allows us to plot the higher-dimensional quantile breakdown frontier in a plane.

3.3 Counterfactual Policies and Unconditional Effects

We work with the potential outcomes framework. For some unknown functions r_0 and r_1

$$\begin{aligned}Y(0) &= r_0(X,U_0),\\ Y(1) &= r_1(X,U_1),\end{aligned}$$

where X are observed covariates and U_0 and U_1 consist of unobservables. We do not impose any restriction on the dimension of the unobservables. The observed outcome is thus

$$\begin{aligned}Y &= D\cdot r_1(X,U_1)+(1-D)\cdot r_0(X,U_0).\\ &:= r(D,X,U),\end{aligned}$$

for a general nonseparable function r, where D is a binary random variable taking values 0 and 1, and $U := (U_0, U_1)'$. The variable D is interpreted as the treatment status, and $p := \Pr(D = 1)$ is the proportion of treated individuals.

In the rest of the chapter, we still maintain Assumption 1.1: a continuity assumption about the outcome Y. This is not essential to our results, but allows us to reduce the notational burden.

A counterfactual policy is an alternative assignment of individuals to treatment. It is given by a binary random variable D_δ, such that $\Pr(D_\delta = 1) = p + \delta$ for a fixed $\delta \in (-p, 1-p)$. It is called counterfactual because it may assign $D_\delta = 1$ to an individual whose $D = 0$. As δ varies over $(-p, 1-p)$, we obtain a collection of counterfactual policies which is denoted by $\mathscr{D}$. Somewhat casually, we also call the collection $\mathscr{D}$ a sequence of policies. When a particular counterfactual policy D_δ belongs to $\mathscr{D}$ we write $D_\delta \in \mathscr{D}$. The counterfactual outcome we would observe for a given $D_\delta \in \mathscr{D}$ is

$$Y_{D_\delta} = r(D_\delta, X, U),$$

where we implicitly assume that the potential outcomes are not affected by the manipulation of D.

We will restrict ourselves to policies that shift a portion of individuals in the control group to the treatment group. We refer to such individuals as *newly treated*. This means that for every individual, $D_\delta - D \geq 0$. This is shown in Figure 3.1.

Assumption 3.1 (Counterfactual Policies)**.** *The sequence of policies $\mathscr{D}$ satisfies*

1. *Monotonicity: $D_\delta - D \geq 0$;*

2. $\Pr(D_\delta = 1) = p + \delta$*;*

3. *The counterfactual outcomes Y_{D_δ} are continuous with positive density on their support $\mathscr{Y}$.*

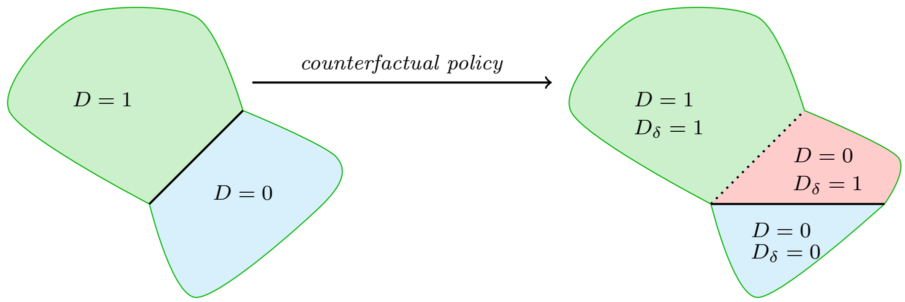

Figure 3.1. A monotonic counterfactual policy.

The monotonicity assumption $D_\delta - D \geq 0$ is mainly for expositional simplicity. Note that it implies that $\delta > 0$. We can do without this assumption, but we need to make some minor changes to our approach. However, there is also a practical purpose. In a context where D is union status, and $D = 1$ denotes unionized individuals, Assumption 3.1 requires that we increase the unionization rate by unionizing previously nonunionized workers. It would probably be hard to simultaneously unionize and deunionize different workers.

Another way to look at the monotonicity assumption is by inspecting the joint distribution of D and D_δ it induces.

Table 3.1. Monotonicity Assumption

	$D_\delta = 0$	$D_\delta = 1$
$D = 0$	$1 - p - \delta$	δ
$D = 1$	0	p

In other words, Assumption 3.1 rules out the presence of *newly untreated* individuals. Also, in the limit, when $\delta = 0$, we return to the original distribution of individuals.

The next task is to define *who* are the *newly treated* individuals, that is, how does D_δ determine who receives treatment among the individuals whose $D = 0$. In this chapter we will focus on two types of policies: a policy that simply chooses individuals whose $D = 0$ at random and assigns them to $D_\delta = 1$, and a policy that chooses individuals based on a user-specified criterion. We will refer to these two types of policies as *randomized policy* and *non-randomized policy* respectively.

Example 3.1 (Randomized policy)**.** *A randomized policy satisfies: for any* $\delta \in [0, 1-p)$

$$D_\delta = \begin{cases} 1 & \text{if } D = 1 \\ 0 \text{ or } 1 & \text{if } D = 0 \end{cases}$$

and the newly treated *are selected at random. Using the conditional independence notation we write* $D_\delta \perp Y(1), Y(0) \| D = 0$*:*

$$Pr(D_\delta = 1 | D = 0) = \Pr(D_\delta = 1 | D = 0, Y(1), Y(0)) = \frac{\delta}{1-p}. \tag{3.1}$$

Example 3.2 (Non-randomized policy)**.** *An example of a non-randomized policy is the following: for any* $\delta \in [0, 1-p)$

$$D_\delta = \begin{cases} 1 & \text{if } D = 1 \\ 1 & \text{if } D = 0 \text{ and } Z \leq F^{-1}_{Z|D=0}\left(\frac{\delta}{1-p}\right) \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

for some observable random variable Z. In this case, the individuals in the group $\{D = 0\}$ *whose Z is less than the* $\frac{\delta}{1-p}$*-quantile of this group are shifted to* $D_\delta = 1$*. This rule guarantees that, in expectation, an aggregate proportion* δ *of individuals is shifted.*

The following theorem characterizes the counterfactual distribution associated with an arbitrary policy.

Theorem 3.1 (Counterfactual Distribution)**.** *For a sequence of policies* $\mathscr{D}$ *that satisfies Assump-*

tions 3.1, the counterfactual distribution for a given $D_\delta \in \mathscr{D}$ is

$$F_{Y_{D_\delta}}(y) = F_a(y) + \delta \left[F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1,D_\delta=1}(y) \right], \tag{3.3}$$

where

$$\begin{aligned} F_a(y) :&= (1-p-\delta) F_{Y(0)|D=0,D_\delta=0}(y) + (p+\delta) F_{Y(1)|D=1,D_\delta=1}(y) \\ &= (1-p-\delta) F_{Y|D=0,D_\delta=0}(y) + (p+\delta) F_{Y|D=1,D_\delta=1}(y). \end{aligned} \tag{3.4}$$

The distribution F_a is called an apparent counterfactual distribution because it is obtained by imputing $F_{Y(1)|D=1,D_\delta=1}$ to the *newly* treated subpopulation.[2] The true distribution, which may not be identifiable, is $F_{Y(1)|D=0,D_\delta=1}$, so the second term corrects this. In a sense, F_a proceeds as if $F_{Y(1)|D=0,D_\delta=1}$ were equal to $F_{Y(1)|D=1,D_\delta=1}$, something which is unlikely to be true. This can be seen in Figure 3.1. The apparent distribution ignores the red shaded area, and combines the green and the blue areas. The second term in (3.3) is the difference between the red and green areas.

The apparent distribution F_a is identified because the policy maker knows the composition of the subpopulations $\{D = 0, D_\delta = 0\}$, the *never* treated, and $\{D = 1, D_\delta = 1\}$, the *already* treated. For both of these subpopulations we observe the "correct" potential outcome. More specifically, for the *never* treated subpopulation, we observe $Y(0)$, and for the *already* treated subpopulation, we observe $Y(1)$. The distributions of Y, which is equal to $Y(0)$ and $Y(1)$, respectively, for the two subpopulations, are identified. As a result, F_a is identified. It is worth pointing out that, under Assumption 3.1, $F_{Y(1)|D=1,D_\delta=1} = F_{Y(1)|D=1}$, and $F_{Y(0)|D=0,D_\delta=0} = F_{Y(0)|D_\delta=0}$. We maintain the "long" notation in order to emphasize the role of D_δ. Also, while $F_{Y(0)|D=0,D_\delta=0}$ is identified, it may not equal $F_{Y|D=0}$ unless D_δ is randomized.

[2] The correct notation for F_a is F_{a,D_δ}, that is, it should include the policy D_δ. However, to keep the notation simple, we omit this. The reader should bear in mind that for two different sequences of policies the apparent distributions might differ.

The only unidentified term in (3.3) is $F_{Y(1)|D=0,D_\delta=1}$. This is the potential outcome $Y(1)$ for the *newly* treated individuals. However, for this subgroup, we only observe $Y(0)$. As we mentioned before, a consequence of Assumption 3.1 is that we lose point identification of the counterfactual distribution.

Remark 3.1 (Firpo et al. (2009))**.** *The marginal effect $M_{\tau,\mathscr{D}}$ was originally studied by Firpo et al. (2009), where, instead of Assumption 3.1, it is assumed that for $d=0,1$: $F_{Y_{D_\delta}|D_\delta=d} = F_{Y|D=d}$. This yields point identification. See the proof to Corollary 3 of the working paper version Firpo et al. (2007). When both D and D_δ are independent of U and X, then $F_{Y_{D_\delta}|D_\delta=d} = F_{Y|D=d}$ will be satisfied. In this particular case, a policy maker can randomize D_δ so that for a given δ, a fraction $p+\delta$ of individuals is randomly assigned to treatment. However, if we allow for D to be endogenous, and if, as is usually the case, the structural form of endogeneity is unknown, then it may be impossible for the policy maker to design a sequence $\mathscr{D}$, such that for every $D_\delta \in \mathscr{D}$, $F_{Y_{D_\delta}|D_\delta=d}$ "matches" $F_{Y|D=d}$. From the point of view of the policy maker, this is a significant restriction on the types of counterfactual policies they can consider.*

Remark 3.2 (Policy Relevant Treatment Effect)**.** *Heckman and Vytlacil (2001b, 2005) and Carneiro et al. (2010, 2011) investigate the effect of a policy on the unconditional mean of the outcome. Using our notation, the Policy Relevant Treatment Effect (PRTE) of Heckman and Vytlacil (2001b, 2005) is*

$$PRTE_{D_\delta} = \frac{E(Y_{D_\delta}) - E(Y)}{\delta},$$

and taking the limit $\delta \to 0$ yields the Marginal PRTE (MPRTE) of Carneiro et al. (2010, 2011):

$$\text{MPRTE}_{\mathscr{D}} = \lim_{\delta \to 0} PRTE_\delta.$$

Martinez-Iriarte and Sun (2020) show how to generalize the MPRTE to cover the case of Firpo et al. (2009) as well.

Remark 3.3 (Rothe (2012))**.** *This paper also studies the global and marginal effects but under* *a different identifying assumption, namely a form of conditional exogeneity. This assumption also yields an identified set. Let the outcome be $Y = r(D, X, U)$. For uniformly distributed random variables $\tilde{U}_1$ and $\tilde{U}_2$, the outcome can be represented as $Y = r(Q_D(\tilde{U}_1), Q_X(\tilde{U}_2), U)$ where Q_D and Q_X are the quantile functions. Then Q_D is changed to another quantile function Q_D^*, generating a counterfactual distribution, which is identified when $\tilde{U}_1 \perp U \| X$ and D is continuous. When D is discrete, $\tilde{U}_1$ is not uniquely determined, so that a range of counterfactual distributions is possible, resulting in partial identification.*

3.3.1 Global Effect: Bounds and Identification Region

We define the *policy selection bias* as the difference between F_{Y_δ} and F_a. We denote it by $\text{psb}(y)$. Since different policies can induce different newly treated individuals, they can induce different counterfactual distributions. Thus the selection bias is policy dependent, and hence the name "policy selection bias."

$$\begin{aligned}\text{psb}(y) :&= \underbrace{F_{Y(1)|D=0,D_\delta=1}(y)}_{\textit{newly}\text{ treated}} - \underbrace{F_{Y(1)|D=1,D_\delta=1}(y)}_{\textit{already}\text{ treated}} \\ &= F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1}(y),\end{aligned}$$

where the second line follows from Assumption 3.1: the subpopulations characterized by $\{D = 1\}$and $\{D = 1, D_\delta = 1\}$ are identical. Under a non-randomized policy, the random variable D_δ is usually a function of D and other observables as in (3.2). So, even if $D \perp Y(0), Y(1)$, that would lead us to

$$\text{psb}(y) = F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=0}(y),$$

which is not zero unless $D_\delta \perp Y(1) \| D = 0$. For example, if we want to choose individuals whose observed outcome is below a certain threshold, then most likely D_δ will be correlated with $Y(1)$ conditional on $D = 0$. Indeed, the fact that $\text{psb}(y)$ is unlikely to be zero seems to be an inevitable feature of the problem.

As a measure of departure from point identification, we will bound the the policy selection bias both from below and from above. For simplicity, we denote by $\mathscr{Y}$ the common support of $Y(1)|D=0,D_\delta=1$ and $Y(1)|D=1,D_\delta=1$.

Assumption 3.2 (*L-U* Bounds)**.** *For any $D_\delta \in \mathscr{D}$, there exists a pair of real numbers $L \in [-1,0]$ and $U \in [0,1]$ such that for every $y \in \mathscr{Y}$*

$$L \leq F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1,D_\delta=1}(y) \leq U,$$

and either $L \neq -1$ or $U \neq 1$.

The distribution of *newly* treated individuals is $F_{Y(1)|D=0,D_\delta=1}(y)$, while $F_{Y(1)|D=1,D_\delta=1}(y)$ is the distribution of the *already* treated individuals. A more precise way to define L is as the infimum over the differences $F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1,D_\delta=1}(y)$, while U is the supremum over such differences.

If $F_{Y(1)|D=1,D_\delta=1}(y)$ first-order stochastically dominates $F_{Y(1)|D=0,D_\delta=1}(y)$, the we can set $L=0$ and $U \leq 1$. In the more general case, where the two distributions cross each other, then $L \in [-1,0]$ and $U \in [0,1]$, and we do not necessarily need to have $U=-L$. Finally, setting $L=-1$ and $U=1$ corresponds to a trivial bounds situation. Since this is always true, this case is excluded from the assumption.

Assumption 3.2 implies via (3.3) that the discrepancy between the counterfactual distribution $F_{Y_{D_\delta}}$ and the apparent distribution F_a is further shrunk by a factor of δ:

$$\delta L \leq F_{Y_{D_\delta}}(y) - F_a(y) \leq \delta U. \tag{3.5}$$

We are now ready to state the main result of this section.

Theorem 3.2 (Global Effect Bounds)**.** *For a given sequence of policies that satisfies Assumptions*

1.1, 3.1 and 3.2, the global effect is bounded by

$$G_{\tau,D_\delta} \in \left[F_a^{-1}(\tau - \delta U) - F_Y^{-1}(\tau), F_a^{-1}(\tau - \delta L) - F_Y^{-1}(\tau)\right] \tag{3.6}$$

for any $\tau \in (\delta U, 1 + \delta L)$.

The proof is best given by a picture. Figure 3.2 shows $F_a(y)$ in solid blue, along with the uniform bounds for $F_{Y_{D_\delta}}(y)$ for given values of L and U. For a fixed τ, $F_{Y_{D_\delta}}^{-1}(\tau)$ must lie between the points ℓ and u. The point ℓ satisfies $F_a(\ell) + \delta U = \tau$, from which we obtain $\ell = F_a^{-1}(\tau - \delta U)$. A similar reasoning applied to u yields $u = F_a^{-1}(\tau - \delta L)$. Finally, the bound for the global effect is obtained by subtracting $F_Y^{-1}(\tau)$ from both ℓ and u.

The identified region in (3.6) is obtained by correcting the evaluation point of the quantile of the apparent distribution: instead of τ, we evaluate the quantile of the apparent distribution at $\tau - \delta U$ and $\tau - \delta L$. The farther away are U and L from zero, where point identification holds, the bigger is the region where the counterfactual distribution can lie. This is reflected in the widening of the identified region.

An important quantity that we will use later on is the *apparent global effect*. This is the estimand that neglects the policy selection bias by setting $L = U = 0$. It is given by

$$G^a_{\tau,D_\delta} := F_a^{-1}(\tau) - F_Y^{-1}(\tau), \tag{3.7}$$

where the superscript "a" conveys the fact that it captures an apparent effect. Indeed, G^a_{τ,D_δ} proceeds as if the counterfactual distribution $F_{Y_{D_\delta}}$ equals the apparent distribution F_a.

The bounds are "monotone" in L and U as we move away from point identification, that is, when $L = U = 0$. Indeed, as we move in the L direction towards -1, the upper bound $F_a^{-1}(\tau - \delta L) - F_Y^{-1}(\tau)$ increases. As we move away from point identification in the U direction towards 1, the lower bound $F_a^{-1}(\tau - \delta U) - F_Y^{-1}(\tau)$ decreases. It is important to recall that L is non-positive, and U is non-negative, so that $F_a^{-1}(\tau - \delta U) \leq F_a^{-1}(\tau - \delta L)$.

Remark 3.4 (Range of τ)**.** *The requirement that* $\tau \in (\delta U, 1+\delta L)$ *comes from* $\tau - \delta U \geq 0$ *and* $\tau - \delta L \leq 1$. *However, later on, when we fix* τ *and* δ, *we want* U *and* L *to not be restricted,* i.e.*, they both can achieve* 1 *and* -1 *respectively. In order for this to happen, we need* $\delta < \tau < 1-\delta$. *In the empirical application we work with* $\delta = 0.1$, *so there will not be a significant restriction on the quantiles we can analyze.*

Remark 3.5 (Trivial Bounds)**.** *In principle, the global effect need not be restricted. The trivial bounds,* $U = 1$ *and* $L = -1$, *provide a bounded region which contains the global effect:*

$$F_a^{-1}(\tau-\delta) - F_Y^{-1}(\tau) \leq G_{\tau,D_\delta} \leq F_a^{-1}(\tau+\delta) - F_Y^{-1}(\tau),$$

so, in the language of Manski (1989, 1990), the trivial bounds are always informative. However, the identified set derived from the trivial bounds always contains 0 *for all quantiles. The intuition is that for a given* δ, *and the common support assumption, we know the counterfactual distributions for a proportion* $1-\delta$ *of individuals. Thus, we are able to bound the quantiles. See the Apprendix for a proof.*

Remark 3.6 (c-dependence)**.** *A common way to relax* $D \perp Y(1)$ *is a version*[3] *of the c-dependence approach of Masten and Poirier (2018), which posits a* $c \in [0,1]$ *such that*

$$\sup_{y \in supp(Y(1))} |\Pr(D=1|Y(1)=y) - \Pr(D=1)| \leq c. \tag{3.8}$$

When $c = 0$, *then* $D \perp Y(1)$. *If* $c > 0$, *then some sort of dependence is allowed between* D *and* $Y(1)$. *Alas, this approach would only help us in the case of randomized policies. For non-randomized policies, to achieve point identification we need an extra conditional independence assumption, namely* $D_\delta \perp Y(1)|D = 0$. *We could, in addition to the c-dependence condition in*

[3]We do not follow exactly the definition of c-dependence of Masten and Poirier (2018) which includes covariates.

(3.8)*, impose*

$$\sup_{y \in supp(Y(1))} |\Pr(D_\delta = 1 | Y(1) = y, D = 0) - \Pr(D_\delta = 1 | D = 0)| \leq c^*.$$

for some $c^* \in [0,1]$. *However, the drawback is that the relationship between c and* c^* *is not at all clear. More importantly, their interpretation is not straightforward either.*

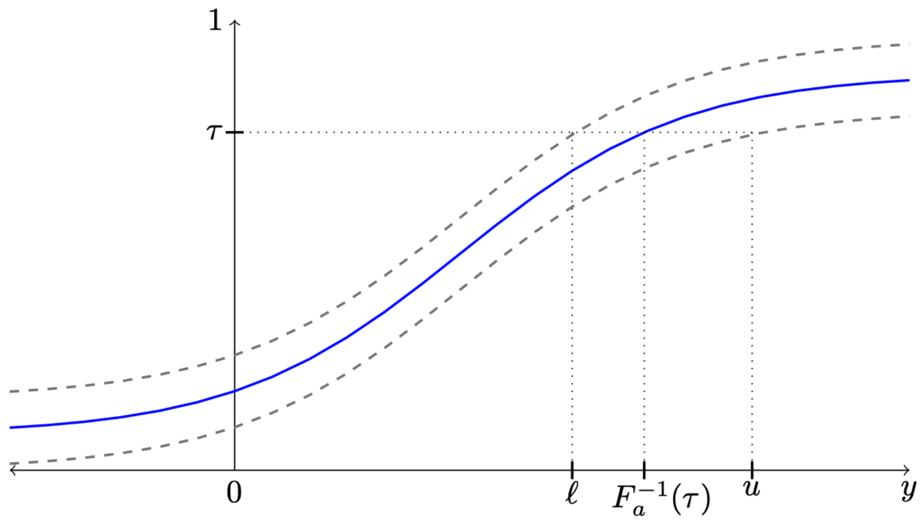

Figure 3.2. Bounds on counterfactual quantiles.

3.3.2 Marginal Effect: Bounds and Identification Region

Before we proceed, we recall the conditions of Theorem 1.1, for the marginal effect to exists. The conditions are

1. $F_{Y_{D_0}}(y) = F_Y(y)$ for any $y \in \mathscr{Y}$;
2. The map $\delta \mapsto F_{Y_{D_\delta}}(y)$ is differentiable at $\delta = 0$ uniformly in $y \in \mathscr{Y}$, with derivative $\dot{F}_{Y,\mathscr{D}}(y)$, that is

$$\lim_{\delta \downarrow 0} \sup_{y \in \mathscr{Y}} \left| \frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} - \dot{F}_{Y,\mathscr{D}}(y) \right| = 0;$$

3. The map $y \mapsto \dot{F}_{Y,\mathscr{D}}(y)$ is continuous at $F_Y^{-1}(\tau)$.

In this case, by Theorem 1.1, the marginal effect exists, is given by

$$M_{\tau,\mathscr{D}} = -\frac{\dot{F}_{Y,\mathscr{D}}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))}.$$

As mentioned before, the conditions and the proof of this Theorem come from viewing the marginal effect as a Hadamard derivative. A primitive condition for $F_{Y_{D_0}}(y) = F_Y(y)$ is Assumption 3.1, because it implies the expansion of Theorem 3.1. Setting $\delta = 0$ in (3.3) yields $F_{Y_{D_0}}(y) = F_Y(y)$. It states that for $D_0 \in \mathscr{D}$, the limiting counterfactual distribution F_{D_0} has to match the observed distribution F_Y. Recall Example 1.1, where $F_{Y_{D_0}}(y) \neq F_Y(y)$.

To better understand what the second condition requires, consider the following rearrangement[4] of equation (3.3):

$$\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} = F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(0)|D=0,D_\delta=1}(y).$$

The right hand side is the difference in potential outcomes for the *newly treated*. We require this change to be continuous in δ: small departures from 0 to $\delta > 0$ should not induce large (uniform) changes in the counterfactual distribution $F_{Y_{D_\delta}}$. This is automatically satisfied when the sequence of policies are randomized. The next example shows this.

Example 3.3 (Marginal Effect of Randomized Policy)**.** *For the case of a randomized policy that satisfies Assumption 3.1, by* (3.1) *we can simplify the counterfactual distribution in* (3.3) *to*

$$F_{Y_{D_\delta}}(y) = F_Y(y) + \delta\left[F_{Y(1)|D=0}(y) - F_{Y(0)|D=0}(y)\right],$$

[4]We can write $F_Y(y) = (1-p-\delta)F_{Y(0)|D=0,D_\delta=0}(y) + \delta F_{Y(0)|D=0,D_\delta=1}(y) + pF_{Y(1)|D=1,D_\delta=1}(y)$, and $F_Y(y) = (1-p-\delta)F_{Y(0)|D=0,D_\delta=0}(y) + \delta F_{Y(0)|D=0,D_\delta=1}(y) + pF_{Y(1)|D=1,D_\delta=1}(y)$. Subtracting $F_Y(y)$ to $F_{Y_{D_\delta}}(y)$ we get

$$F_{Y_{D_\delta}}(y) - F_Y(y) = \delta\left(F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(0)|D=0,D_\delta=1}(y)\right).$$

which implies that $\dot{F}_{Y,\mathscr{D}}(y) = F_{Y(1)|D=0}(y) - F_{Y(0)|D=0}(y)$. *We obtain that* $\dot{F}_{Y,\mathscr{D}}$ *is independent of* $\mathscr{D}$.

In the next example, we show how the uniform differentiability might fail with arbitrary non-randomized policies.

Example 3.4 (Effect at the Margin)**.** *For the case of a non-randomized policy that satisfies Assumption 3.1, equation* (3.3) *in Theorem 3.1 can be written as*

$$\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} = F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(0)|D=0,D_\delta=1}(y).$$

It is not immediate that the limit when δ *goes to* 0 *of the right hand side exists pointwise for any* $y \in \mathscr{Y}$. *This is not obvious, since the conditioning set* $D = 0, D_\delta = 1$ *shrinks to a measure* 0 *set: those individuals whose* $D = 0$ *and* $D = 1$.

In the case of a threshold-crossing model for D, as in $D = \mathbb{1}\left\{V \leq F_V^{-1}(p)\right\}$, *and a sequence of policies such that* $D_\delta = \mathbb{1}\left\{V \leq F_V^{-1}(p+\delta)\right\}$, *the event* $\{D = 0, D_\delta = 1\}$ *is equivalent to the event* $\left\{0 \leq V \leq F_V^{-1}(p+\delta)\right\}$, *so we can define the limiting conditioning probability to be*

$$\begin{aligned}\lim_{\delta\to 0}\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} &= \lim_{\delta\to 0} F_{Y(1)|D=0,D_\delta=1}(y) - \lim_{\delta\to 0} F_{Y(0)|D=0,D_\delta=1}(y) \\ &= F_{Y(1)|V=0}(y) - F_{Y(0)|V=0}(y).\end{aligned}$$

If the distributions $F_{Y(1)|V=0}$ *and* $F_{Y(0)|V=0}$ *are continuous then pointwise convergence is equivalent to uniform convergence. This implies that* $\dot{F}_{Y,\mathscr{D}}(y) = F_{Y(1)|V=0}(y) - F_{Y(0)|V=0}(y)$ *which is the familiar result (see Martinez-Iriarte and Sun (2020)) that the marginal individuals, those whose* $V = 0$, *are the ones that drive the marginal effect in a threshold-crossing model. The effect is policy dependent because it entails a particular departure in the selection equation. This policy dependence is emphasized in Carneiro et al. (2010, 2011).*

When $\dot{F}_{Y,\mathscr{D}}$ is independent of $\mathscr{D}$, two different randomized policies $\mathscr{D}$ and $\mathscr{D}'$ will deliver the same marginal effect. But at the same time, at the population level, there can only be one randomized policy. This result reflects precisely that. On the other hand, for non-randomized policies, the marginal effect can easily be sequence dependent.

For a randomized policy, $\dot{F}_{Y,\mathscr{D}}$ is well-defined, though not necessarily identified since it involves $F_{Y(1)|D=0}$. For this reason, when analyzing the marginal effect, we will focus on randomized policies. Thus, we will write M_τ instead of $M_{\tau,\mathscr{D}}$.

The bounds for the marginal effect will be obtained as the limiting bounds, as δ goes to 0, for the global effect under a randomized policy. That is, the bounds for M_τ will be given by

$$\lim_{\delta\to 0}\frac{F_a^{-1}(\tau-\delta U)-F_Y^{-1}(\tau)}{\delta},$$

and

$$\lim_{\delta\to 0}\frac{F_a^{-1}(\tau-\delta L)-F_Y^{-1}(\tau)}{\delta},$$

provided these limits exist.

These limits can be seen as derivatives with respect to δ of $\delta\mapsto F_a^{-1}(\tau-\delta L)$ at $\delta=0$. There is a minor complication which makes the computation a bit more involved. The reason is that δ plays a dual role in the map $\delta\mapsto F_a^{-1}(\tau-\delta L)$: first, it enters in the argument of $F_a^{-1}(\tau-\delta L)$; second it is used in the construction of the apparent distribution $F_a := (1-p-\delta)F_{Y|D=0}+(p+\delta)F_{Y|D=1}$ (see (3.4)). We resort to the chain rule and treat each case separately. The first case can be solved as an ordinary derivative of the inverse of a function, while the second case takes advantage of the Hadamard differentiability of the function $\delta\mapsto F_a$, which maps a scalar into the space of right-continuous functions with left limits, composed with the function $F_a\mapsto F_a^{-1}(\tau)$ which maps an increasing right-continuous function with left limits into the real numbers. The details can be found in the Appendix. Heuristically, we have

$$\lim_{\delta\to 0}\frac{F_a^{-1}(\tau-\delta U)-F_Y^{-1}(\tau)}{\delta}=\lim_{\delta\to 0}\frac{F_Y^{-1}(\tau-\delta U)-F_Y^{-1}(\tau)}{\delta}+\lim_{\delta\to 0}\frac{F_a^{-1}(\tau)-F_Y^{-1}(\tau)}{\delta},$$

where the first term can be dealt with the inverse function theorem, and the second term with a Hadamard derivative to account how the function F_a^{-1} moves when we move δ.

Theorem 3.3 (Marginal Effect Bounds)**.** *For a sequence of randomized policies that satisfies Assumptions 1.1, 3.1 and 3.2 the marginal effect is bounded by*

$$-\frac{U}{f_Y(F_Y^{-1}(\tau))}\leq M_\tau - M_\tau^a \leq -\frac{L}{f_Y(F_Y^{-1}(\tau))} \tag{3.9}$$

for any $\tau\in(0,1)$*, where*

$$M_\tau^a := -\frac{F_{Y|D=1}(F_Y^{-1}(\tau))-F_{Y|D=0}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))} \tag{3.10}$$

is the apparent effect.

The apparent effect in (3.10) is the estimand of Firpo et al. (2009). Hence, Theorem 3.3 states that the usual estimand should be enlarged by $-\frac{U}{f_Y(F_Y^{-1}(\tau))}$ and $-\frac{L}{f_Y(F_Y^{-1}(\tau))}$ in order to contain M_τ. Recall that L is non-positive, and U is non-negative. As opposed to the bounds on the global effect, the result in Theorem 3.3 holds for any $\tau\in(0,1)$. However, there is not much to gain from this because as τ approaches 0 or 1, the density $f_Y(F_Y^{-1}(\tau))$ is likely to approach zero and the bounds will diverge to $+\infty$ or $-\infty$.

Remark 3.7 (Trivial Bounds)**.** *Setting* $L=-1$ *and* $U=1$ *corresponds to a trivial bounds case. It is a matter of simple algebra to show that* 0 *will always be in the identified set in this case. For example, if* $M_\tau^a\geq 0$*, then* $0\in[M_\tau^a-1/f_Y(F_Y^{-1}(\tau)), M_\tau^a+1/f_Y(F_Y^{-1}(\tau))]$*. As in the case with the global effect, the boundedness of the outcome is not needed for the trivial bounds to be informative.*

3.4 Quantile Breakdown Frontier

In our framework, the amount of policy selection bias is controlled by L and U. Figure 3.3 shows this in the $L \times U$ plane. When $L = U = 0$, there is no policy selection bias, and hence we achieve point identification. Any other value of $L \in [-1, 0)$ and $U \in (0, 1]$ admits some policy selection bias, and consequently the effects are only partially identified. A special case of this are the trivial bounds: when $L = -1$ and $U = 1$. We refer to any combination (L, U) distinct from $(0, 0)$ as a *departure from point identification*.

The following language convention is important. Because L is always non-positive, we say that we have *more* policy selection bias (due to L) in the point $(L, U) = (-1, u)$ than in the point $(L, U) = (-0.5, u)$, even though L is bigger in the latter, -0.5, than in the former, -1. Thus, we quantify the selection as how far we move (L, U) from $(0, 0)$, rather than by the value of L or U.

The quantile breakdown frontier is a curve that quantifies the amount of policy selection bias compatible with a given conclusion of interest across quantiles. Suppose we are interested in a certain policy D_δ, and we would like to know if its global effect on the median of Y is positive. That is, we want to know whether $G_{.5,D_\delta} > 0$ or not. If we were certain that there is no policy selection bias, we would just estimate the apparent effect $G^a_{.5,D_\delta}$ using (3.7):

$$G^a_{.5,D_\delta} = F_a^{-1}(.5) - F_Y^{-1}(.5).$$

However, it is very likely that the apparent effect $G^a_{.5,D_\delta}$ is biased for the true global effect $G_{.5,D_\delta}$. Sensitivity analysis, in a sense, asks the reverse question: how much policy selection bias is compatible with $G_{.5,D_\delta} > 0$? The quantile breakdown frontier answers this question by indicating the amount of departure from point identification such that the conclusion holds.

In order to answer the question posed by sensitivity analysis, we recall Theorem 3.2

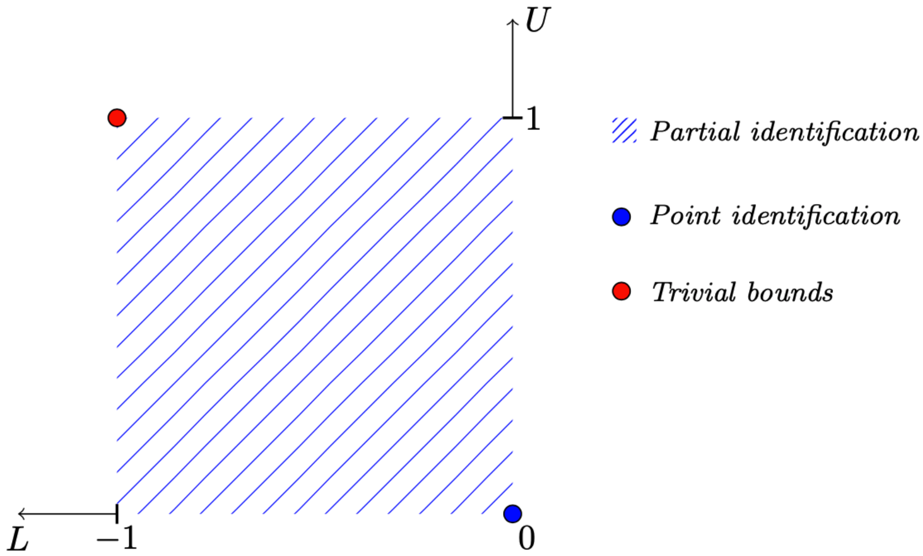

Figure 3.3. Point and partial identification.

which states that there are L and U such that

$$F_a^{-1}(.5-\delta U)-F_Y^{-1}(.5)\leq G_{.5,D_\delta}\leq F_a^{-1}(.5-\delta L)-F_Y^{-1}(.5). \tag{3.11}$$

Hence, for $G_{.5,D_\delta}>0$ to hold, we need that the lower bound in (3.11) be greater than zero. That is, we need all the values of U such that

$$0<F_a^{-1}(.5-\delta U)-F_Y^{-1}(.5)\leq G_{.5,D_\delta} \tag{3.12}$$

First, we note that $F_a^{-1}(.5-\delta U)$ is decreasing in U. Suppose $G_{.5,D_\delta}$¿0. We start with $U=0$ and then move it towards 1. On the other hand, L is left unrestricted. So, all the values of U such that (3.12) holds and any value of $L\in[-1,0]$ are compatible with $G_{.5,D_\delta}>0$. In particular, let $U_{.5}$ be the value of U such that the lower bound in (3.12) is equal to zero: $0=F_a^{-1}(.5-\delta U_{.5})-F_Y^{-1}(.5)$. Thus, the combination of L and U compatible with $G_{.5,D_\delta}>0$

are

$$\{(L,U) : -1 \leq L \leq 0 \text{ and } 0 \leq U < U_{.5}\} \tag{3.13}$$

A value of U greater than $U_{.5}$ induces too much bias and fails to guarantee that $G_{.5,D_\delta} > 0$. The left panel of Figure 3.4 shows the compatible values of L and U in the $L \times U$ plane.

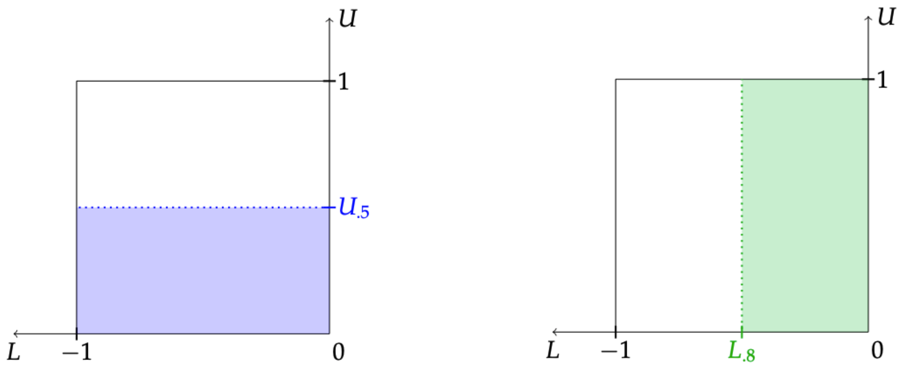

Figure 3.4. Compatible values for $G_{.5,D_\delta} > 0$ and $G_{.8,D_\delta} < 0$.

Now suppose we are also interested in the 80th quantile. However, it may be the case that there is no value of U such that

$$0 < F_a^{-1}(.8 - \delta U) - F_Y^{-1}(.8).$$

holds. That is, for any value of U,

$$F_a^{-1}(.8 - \delta U) - F_Y^{-1}(.8) \leq 0,$$

or equivalently $G^a_{.8,D_\delta} \leq 0$, thus, no combination of L and U can *guarantee* that $G_{.8,D_\delta} > 0$ holds. Therefore, we look at the reverse conclusion $G_{.8,D_\delta} < 0$, and find all the values of L such that

$$G_{.8,D_\delta} \leq F_a^{-1}(.8 - \delta L) - F_Y^{-1}(.8) < 0.$$

We denote by $L_{.8}$ the value of L that solves: $F_a^{-1}(.8-\delta L)-F_Y^{-1}(.8)=0$. The values of L and U such that $G_{.8,D_\delta}<0$ are

$$\{(L,U): L_{.8} < L \leq 0 \text{ and } 0 \leq U \leq 1\}, \tag{3.14}$$

and are shown in the right panel of Figure 3.4.

The extension of this procedure to more than two quantiles gives rise to the quantile breakdown frontier. For a collection of conclusions indexed by $\tau \in (\delta, 1-\delta)$,[5] for example $G_{\tau,D_\delta} > g_\tau$, the quantile breakdown frontier shows the combinations of L and U compatible with each conclusion.

Figure 3.5 contains an hypothetical quantile breakdown frontier constructed for all $\tau \in (\delta, 1-\delta)$. On the left side, at $\tau = .5$, we can see that below the curve we have the region described in (3.13) under which $G_{.5,D_\delta} > 0$ holds. At $\tau = .8$, we have that *above* the curve we have the region described in (3.14) where $G_{.8,D_\delta} < 0$ holds. The right hand side shows this for all the quantiles in $(\delta, 1-\delta)$. Values of U in the red area include possible negative values of the global effect. The green area is the counterpart of the blue area: a robust region for $G_{\tau,D_\delta} < 0$. Finally, the orange area is the counterpart of the red area: values of L such that the global effect might be positive.

Consider again the left panel in Figure 3.5. We can use the values $L_{.8}$ and $U_{.5}$ to construct bounds for the global effect curve: $\tau \mapsto G_{\tau,D_\delta}$. These bounds have the property that at $\tau = 0.5$, the identified region for the global effect is positive, while at $\tau = 0.8$, the identified region for the global effect is negative. Moreover, the identified region for the global effect derived from $L_{.8}$ and $U_{.5}$ will provide statements about the global effect at other quantiles as well. This can be seen in Figure 3.6. The solid line in upper left panel shows the trivial bounds. These are obtained by setting $L = -1$ and $U = 1$ in (3.6). Note how the identified region of the trivial bounds contains 0 for all the quantiles, in line with Remark 3.5. The upper right panel shows

[5] See Remark 3.4 for an explanation of this restriction.

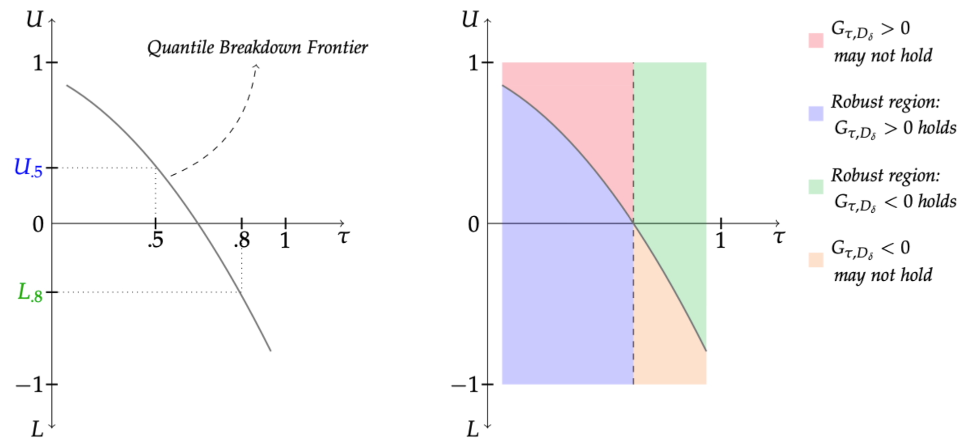

Figure 3.5. Quantile Breakdown Frontier.

the restriction on U such that $G_{.5,D_\delta} > 0$: the lower bounds is tightened and crosses 0 at $\tau = 0.5$. Similarly, the lower left panel tightens the upper bound consistent with restricting L to be $L_{.8}$ in order for $G_{.8,D_\delta} < 0$ to hold. Note how the upper bound now crosses 0 at $\tau = 0.8$. The lower right panel gives simultaneous bounds for the global effect such that $G_{.5,D_\delta} > 0$ and $G_{.8,D_\delta} < 0$. The interpretation of the grey shaded area in the lower right panel of Figure 3.6 is the following: the global effect curve has to lie in the gray area in order for the conclusions $G_{.5,D_\delta} > 0$ and $G_{.8,D_\delta} < 0$ to hold.

One of the building blocks for the construction of the quantile breakdown frontier is the breakdown frontier of Masten and Poirier (2020). Figure 3.4 shows two examples a breakdown frontier. The quantile breakdown frontier takes advantage of the fact that the frontiers in Figures 3.4 are straight lines. This simplicity allows us to plot the higher dimensional quantile breakdown frontier in a plane as in Figure 3.5.

In the rest of this section we will derive analytical expressions for L_τ, U_τ, the quantile breakdown frontier, and the bounds on the global effect. We will also derive the quantile breakdown frontier for the sign of the marginal effect.

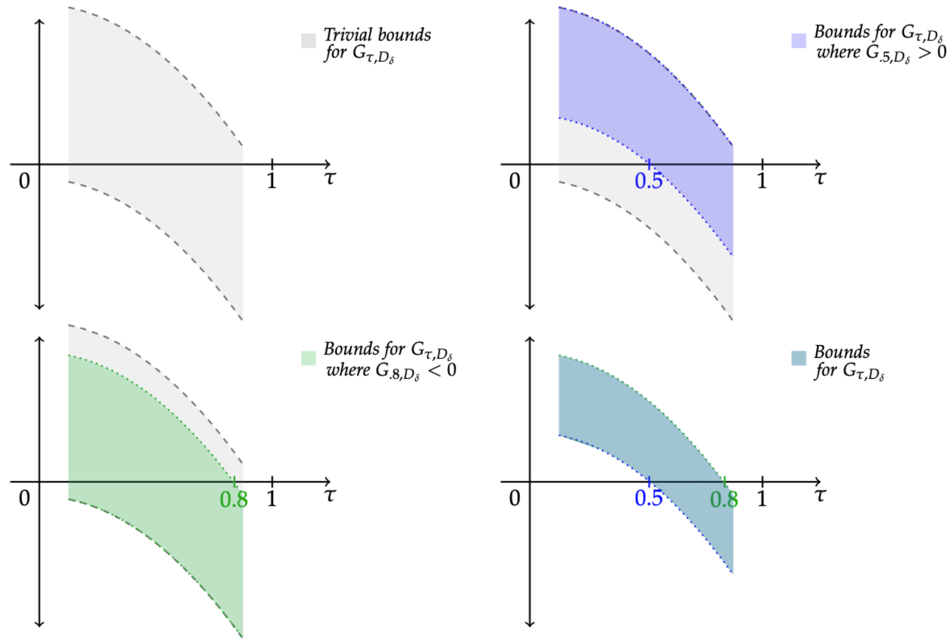

Figure 3.6. Bounds on the global effect.

3.4.1 Global Effect

Suppose that, for a given τ and D_δ, we are interested in the global effect. By Theorem 3.2, there are L and U such that

$$F_a^{-1}(\tau - \delta U) - F_Y^{-1}(\tau) \leq G_{\tau,D_\delta} \leq F_a^{-1}(\tau - \delta L) - F_Y^{-1}(\tau).$$

In order not to impose restrictions of L and U, we will focus on $\tau \in (\delta, 1-\delta)$ (See Remark 3.4). We further recall that the bounds are "centered" around $G^a_{\tau,D_\delta} := F_a^{-1}(\tau) - F_Y^{-1}(\tau)$. For a given τ we are interested in the values of L and U such that either $G_{\tau,D_\delta} > g_\tau$ or $G_{\tau,D_\delta} < g_\tau$ holds. In order to build the breakdown frontier we must look at the location of G^a_{τ,D_δ} with respect to g_τ.

Figure 3.7 illustrates the case of $G^a_{\tau,D_\delta} > g_\tau$. The blue part of the axis shows the possible values of G_{τ,D_δ}. The dashed lines show three different combination of L and U. The two blue dashed lines allow us to conclude that the effect is greater than g_τ. The red dashed line include values lower than g_τ, and hence it is excluded. Since only the lower bounds concern us, this means that we do not want U to get too close to 1. Thus there is a maximum departure from point identification due to U that ensures that $G_{\tau,D_\delta} > g_\tau$ holds in the case where $G^a_{\tau,D_\delta} > g_\tau$.

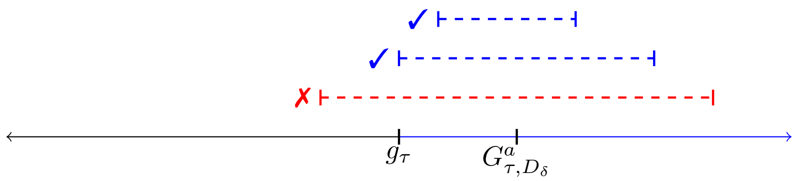

Figure 3.7. Maximum departure from $U = 0$.

This maximum U is denoted by U_τ, and it solves (see middle dashed line in Figure 3.7)

$$F_a^{-1}(\tau - \delta U_\tau) - F_Y^{-1}(\tau) = g_\tau.$$

which implies that

$$U_\tau = \min\left\{\max\left\{0, \frac{\tau - F_a(F_Y^{-1}(\tau) + g_\tau)}{\delta}\right\}, 1\right\}. \tag{3.15}$$

Figure 3.8 shows the other possibility, which is $G^a_{\tau,D_\delta} < g_\tau$. In this case we can analyze conclusions of the form $G_{\tau,D_\delta} < g_\tau$. As we move L towards -1, the right end of the identified regions approaches g_τ. In the red segment, L is too close to -1, so the identified region contains values contrary to the conclusion. So, it is excluded.

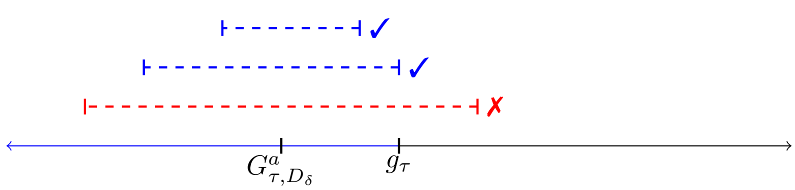

Figure 3.8. Maximum departure from $L = 0$.

In this case, we have

$$L_\tau = \max\left\{\min\left\{0, \frac{\tau - F_a(F_Y^{-1}(\tau) + g_\tau)}{\delta}\right\}, -1\right\}. \tag{3.16}$$

The common ingredient for U_τ in (3.15) and L_τ in (3.16) is

$$\theta(\tau) = \frac{\tau - F_a(F_Y^{-1}(\tau) + g_\tau)}{\delta}. \tag{3.17}$$

The map $\tau \mapsto \theta(\tau)$ is the quantile breakdown frontier. Alternatively, for a given τ, the quantile breakdown frontier is value of the policy selection bias such that the global effect $G_{\tau,D_\delta} = g_\tau$. If this value is positive, it is taken to be U, if it is negative, it is taken to be L.

Continuity of the quantile breakdown frontier is important for inference purposes. Inspection of the formulas in (3.17) shows that continuity of F_a, F_Y^{-1}, and of the map $\tau \mapsto g_\tau$ is enough. Continuity of F_a and F_Y^{-1} is true by assumption, but continuity of $\tau \mapsto g_\tau$ is up to the user. In our empirical application we will choose a g_τ which is constant-across-τ. An arbitrary collection of $\{g_\tau : \tau \in (\delta, 1-\delta)\}$ might be problematic.

Summarizing, we follow the following steps. First, we need to fix the set of quantiles τ in which we are interested and compute the quantile breakdown frontier for a given collection of g_τ. Then, we have to check the sign of the quantile breakdown frontier at these τ's. If the quantile breakdown frontier is positive, we can derive the values of U such that positive conclusions hold: $G_{\tau,D_\delta} > g_\tau$. If the quantile frontier is negative, we can derive the values of L such that the negative conclusions hold: $G_{\tau,D_\delta} < g_\tau$.

3.4.2 Bounds derived from the QBF

Often times, researchers are interested in quantile contrasts: for example Farber et al. (2020) examine 10^{th} vs. 90^{th} of a marginal increase in unionization. In this case, following Masten and Poirier (2020) we can visualize the result in a joint breakdown frontier/robust region.

This is shown in Figure 3.9. The intersection contains the values of L and U compatible with both conclusions of interest. We can use the quantile breakdown frontier to derive bounds on the

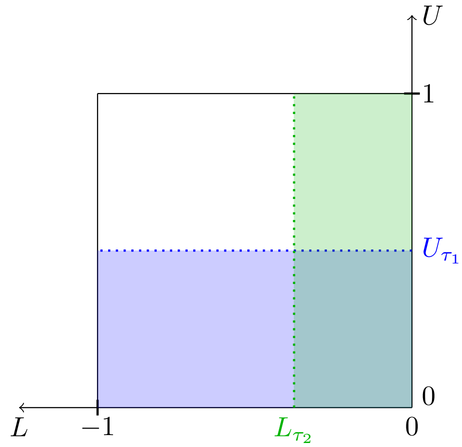

Figure 3.9. Joint Breakdown Frontier.

global effect for every $\tau \in (\delta, 1-\delta)$:

$$\tau \mapsto G_{\tau, D_\delta}.$$

To do so, we find τ_1 and τ_2 such that we can analyze $G_{\tau_1} > 0$ and $G_{\tau_2} < 0$. That is, the quantile breakdown is positive at τ_1 and negative at τ_2.[6] Following Theorem 3.2, we can use U_{τ_1} to construct a lower bound for the global effect, and L_{τ_2} to construct an upper bound for the global effect. These bounds are given by

$$\tau \mapsto B(U_{\tau_1}; \tau) := F_a^{-1}(\tau - \delta U_{\tau_1}) - F_Y^{-1}(\tau), \tag{3.18}$$

and

$$\tau \mapsto B(L_{\tau_2}; \tau) := F_a^{-1}(\tau - \delta L_{\tau_2}) - F_Y^{-1}(\tau). \tag{3.19}$$

[6] The empirical quantile breakdown frontier might be negative or positive everywhere. In that case this analysis would not apply. However, in our empirical analysis, the quantile breakdown is positive in a region, and negative in another region.

and are shown in Figure 3.6.

3.4.3 Marginal Effect

For the marginal effect the situation is a bit more delicate, and some care must be exercised with the density in the denominator. By Theorem 3.3, the identified region for M_τ is

$$\left[M_\tau^a - \frac{U}{f_Y(F_Y^{-1}(\tau))}, M_\tau^a - \frac{L}{f_Y(F_Y^{-1}(\tau))}\right]$$

where recall that

$$M_\tau^a := -\frac{F_{Y|D=1}(F_Y^{-1}(\tau)) - F_{Y|D=0}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))}.$$

Consider a single quantile τ. For the conclusion $M_\tau > g_\tau$ to hold, then, as before, the restriction on U, denoted by U_τ, solves

$$-\frac{F_{Y|D=1}(F_Y^{-1}(\tau)) - F_{Y|D=0}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))} - \frac{U_\tau}{f_Y(F_Y^{-1}(\tau))} = g_\tau$$

which implies

$$U_\tau = \min\left\{1, \max\left\{0, F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau)) - g_\tau f_Y(F_Y^{-1}(\tau))\right\}\right\}.$$

For the opposite conclusion, $M_\tau < g_\tau$, similar calculations, this time on the upper bound, yield

$$L_\tau = \max\left\{\min\left\{0, F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau)) - g_\tau f_Y(F_Y^{-1}(\tau))\right\}, -1\right\}.$$

When it comes to estimation, the quantile breakdown frontier contains a non-parametric ingredient, namely the density f_Y evaluated at a quantity that must estimated: $F_Y^{-1}(\tau)$. This can be avoided if we set $g_\tau = 0$ for every τ. In such a case, we are interested in the sign of the marginal effect. This is natural conclusion to be interested in since the marginal effect has the interpretation of a derivative. When $g_\tau = 0$, these expressions simplify to

$$U_\tau = \min\left\{\max\left\{0, F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau))\right\}, 1\right\},$$

and

$$L_\tau = \max\left\{\min\left\{0, F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau))\right\}, -1\right\}.$$

The quantile breakdown frontier for the sign of marginal effect is then given by

$$\tau \mapsto \theta(\tau) = F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau)). \tag{3.20}$$

Remark 3.8. *Coincidentally, in this case where $g = 0$ for every τ, the quantile breakdown frontiers for the global and the marginal effects coincide. This reflects the fact that the apparent marginal effect and the apparent global effects have the same sign. Of course, the true effects might differ in sign. To see this, we note that the apparent distribution can be written as*[7] $F_a(y) = F(y) + \delta\left[F_{Y|D=1,D_\delta=1}(y) - F_{Y|D=0,D_\delta=1}(y)\right]$. *So, that for $g = 0$, and plugging the previous expression for $F_a(y)$ in* (3.17)*, we obtain* (3.20)*.*

3.5 Estimation and Inference

There are two main results in the sensitivity analysis we propose. The first one is the quantile breakdown frontier $\tau \mapsto \theta(\tau)$. The second important result is the case when we use the estimated values of U_{τ_1} and L_{τ_2} to construct bounds for the effect across all quantiles in the manner of Figure 3.6. We will provide asymptotic results both pointwise, for a given τ, and uniform, when the objects are seen as a random function.

We work in the space $\ell^\infty(\delta, 1-\delta)$ of bounded real-valued functions defined on $(\delta, 1-\delta)$.

[7] See the proof for the statement of Remark 3.5.

As usual, we endow this space with the supremum norm: $\|x\|_\infty := \sup_{t\in(\delta,1-\delta)}|x(t)|$. The reason we restrict the space to be $\ell^\infty(\delta,1-\delta)$ and not $\ell^\infty(0,1)$ is due to the fact that for a given δ, we cannot reach quantiles below δ or above $1-\delta$. See Remark 3.4 above.

In order to simplify notation, and ensure the continuity of the quantile breakdown frontier, we are going to focus on the case where the threshold g_τ is constant across τ.

Assumption 3.3 (Constant Threshold)**.** *For some scalar g, the threshold g_τ satisfies $g_\tau = g$ for any $\tau \in (\delta, 1-\delta)$.*

This assumption can be relaxed at the expense of more complicated notation. However, we still require smoothness in the map $\tau \mapsto g_\tau$. For the case of the quantile breakdown for the sign of the marginal effect, we will set $g = 0$.

3.5.1 Quantile Breakdown Frontier: Global Effect

Under Assumption 3.3, the quantile breakdown frontier is

$$\theta(\tau) := \frac{\tau - F_a(F_Y^{-1}(\tau)+g)}{\delta}.$$

The empirical apparent distribution is

$$\hat{F}_a(y) = (1-\hat{p}-\delta)\hat{F}_{Y|D=0,D_\delta=0}(y) + (\hat{p}+\delta)\hat{F}_{Y|D=1,D_\delta=1}(y),$$

where $\hat{p} := n^{-1}\sum_{i=1}^n D_i$, and

$$\hat{F}_{Y|D=0,D_\delta=0}(y) := \frac{\sum_{i=1}^n \mathbb{1}\{Y_i \le y\}(1-D_i)(1-D_{\delta,i})}{\sum_{i=1}^n (1-D_i)(1-D_{\delta,i})},$$
$$\hat{F}_{Y|D=1,D_\delta=1}(y) := \frac{\sum_{i=1}^n \mathbb{1}\{Y_i \le y\} D_i D_{\delta,i}}{\sum_{i=1}^n D_i D_{\delta,i}}.$$

The empirical quantiles $\hat{F}_Y^{-1}$, are computed using the generalized inverse: $\hat{F}_Y^{-1}(\tau) := \inf\left\{y : \hat{F}_Y(y) \ge \tau\right\}$. Here, $\hat{F}_Y(y) := n^{-1}\sum_{i=1}^n \mathbb{1}\{Y_i \le y\}$ is the empirical CDF. For given τ, g and

δ, the estimated counterpart of $\theta(\tau)$ is then

$$\hat{\theta}(\tau) := \frac{\tau - \hat{F}_a(\hat{F}_Y^{-1}(\tau) + g)}{\delta}. \tag{3.21}$$

We can view the map $\tau \mapsto \hat{\theta}(\tau)$ as a random element of $\ell^\infty(\delta, 1-\delta)$. In that case, we denote it simply by $\hat{\theta}$. We want to investigate the weak convergence of $\sqrt{n}(\hat{\theta} - \theta)$ in $\ell^\infty(\delta, 1-\delta)$:

$$\sqrt{n}(\hat{\theta} - \theta) = -\frac{1}{\delta}\sqrt{n}\left(\hat{F}_a \circ (\hat{F}_Y^{-1} + g) - F_a \circ (F_Y^{-1} + g)\right).$$

This is similar to a quantile-quantile transformation (see Exercise 4 in Chapter 3.9 in van der Vaart and Wellner (1996)). We base our proof of the asymptotic distribution of $\sqrt{n}(\hat{\theta} - \theta)$ on the proof of Lemma A.1 in Beare and Shi (2019).[8] The main assumption is

Assumption 3.4 (Functional CLT)**.** *The following multivariate functional central limit theorem holds*

$$\sqrt{n}\begin{pmatrix} \hat{F}_Y - F_Y \\ \hat{F}_{Y|D=0,D_\delta=0} - F_{Y|D=0,D_\delta=0} \\ \hat{F}_{Y|D=1,D_\delta=1} - F_{Y|D=1,D_\delta=1} \\ \hat{p} - p \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{G}_Y \\ \mathbb{G}_{0,0} \\ \mathbb{G}_{1,1} \\ \mathbb{Z}_p \end{pmatrix},$$

where $\mathbb{G}_Y$, $\mathbb{G}_{0,0}$, and $\mathbb{G}_{1,1}$ are Brownian bridges in $\ell^\infty(\mathscr{Y})$, and $\mathbb{Z}_p$ is a (real-valued) normal random variable.

The following assumption is needed to establish the Hadamard differentiable of different functions used in the construction of θ.

Assumption 3.5 (Conditions for Hadamard Differentiability)**.**

[8]Beare and Shi (2019) also offer some interesting historical context for the result.

1. *For some $\varepsilon > 0$, F_Y is continuously differentiable in $[F_Y^{-1}(\delta) - \varepsilon, F_Y^{-1}(1-\delta) + \varepsilon] \subset \mathscr{Y}$ with strictly positive derivative f_Y.*

2. *The distribution functions $F_{Y|D=0,D_\delta=0}(y)$ and $F_{Y|D=1,D_\delta=1}(y)$ are differentiable, with uniformly continuous and bounded derivatives on their support $\mathscr{Y}$.*

The first item in Assumption 3.5 concerns the support $\mathscr{Y}$ and the smoothness of F_Y. It is used to guarantee the Hadamard differentiability of the quantile process $\tau \mapsto F_Y^{-1}(\tau)$ for $\tau \in (\delta, 1-\delta)$. The second item ensures that the apparent distribution $F_a(y)$ has a uniformly continuous and bounded derivative. This derivative is denoted by $f_a(y)$. It is needed to establish the Hadamard differentiability of the composition map $(F_a, F_Y^{-1}) \mapsto F_a \circ (F_Y^{-1} + g)$.[9]

Theorem 3.4 (Asymptotic Distribution of QBF)**.** *Under Assumptions 3.3, 3.4, and 3.5*

$$\sqrt{n}(\hat{F}_a - F_a) \rightsquigarrow \mathbb{G}_a := (1-\delta)\mathbb{G}_{0,0} + \delta\mathbb{G}_{1,1} + (F_{Y|D=1,D_\delta=1} - F_{Y|D=0,D_\delta=0})\mathbb{Z}_p,$$

where $\mathbb{G}_a$ is a Gaussian tight element of $\ell^\infty(\mathscr{Y})$, and

$$\sqrt{n}(\hat{\theta} - \theta) \rightsquigarrow \mathbb{G}_\theta := -\frac{1}{\delta}\mathbb{G}_a \circ (F_Y^{-1} + g) + \frac{1}{\delta} f_a \circ (F_Y^{-1} + g)\frac{\mathbb{G}_Y \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}},$$

where $\mathbb{G}_\theta$ is Gaussian tight element of $\ell^\infty(\delta, 1-\delta)$.

The second convergence result of Theorem 3.4 is uniform in $\tau \in (\delta, 1-\delta)$. If we are interested in a particular quantile τ, we can evaluate $\sqrt{n}(\hat{\theta} - \theta)$ at τ to obtain

$$\begin{aligned}\sqrt{n}(\hat{\theta}(\tau) - \theta(\tau)) \rightsquigarrow \mathbb{G}_\theta(\tau) = &-\frac{1}{\delta}\mathbb{G}_a \circ (F_Y^{-1}(\tau) + g) \\ &+ \frac{1}{\delta} f_a \circ (F_Y^{-1}(\tau) + g)\frac{\mathbb{G}_Y \circ F_Y^{-1}(\tau)}{f_Y \circ F_Y^{-1}(\tau)}.\end{aligned}$$

Instead of providing a closed form expression and a consistent estimator for the variance

[9] Section 3.9 in van der Vaart and Wellner (1996) studies the Hadamard differentiability of composition maps.

of $\mathbb{G}_\theta(\tau)$, we note that, by Theorem 23.9 in van der Vaart (1998), the empirical bootstrap is valid. Confidence intervals for $\theta(\tau)$ can be constructed in the following way:

Algorithm 1 (Bootstrap for $\theta(\tau)$)**.**

1. *Given the data* $\{Y_i, D_i, D_{\delta,i}\}_{i=1}^n$ *and a value* $\tau \in (\delta, 1-\delta)$*, compute* $\hat{\theta}(\tau)$ *as in* (3.21)*.*

2. *Obtain B bootstrap samples of size n from* $\{Y_i, D_i, D_{\delta,i}\}_{i=1}^n$*, and compute* $\sqrt{n}(\hat{\theta}^b(\tau) - \hat{\theta}(\tau))$*, where* $\hat{\theta}^b(\tau)$ *is computed as in* (3.21) *for* $b = 1, \ldots, B$*.*

3. *For* $\left\{\sqrt{n}(\hat{\theta}^b(\tau) - \hat{\theta}(\tau))\right\}_{b=1}^B$*, obtain the* $(100 \times \alpha/2)\%$ *and* $(100 \times (1-\alpha/2))\%$ *percentiles. These are denoted* $\xi_{\alpha/2,\theta(\tau)}$ *and* $\xi_{1-\alpha/2,\theta(\tau)}$*.*

The $1-\alpha$ confidence intervals are then computed as

$$\mathscr{CI}(\theta(\tau), \alpha) = \left[\hat{\theta}(\tau) - \frac{\xi_{1-\alpha/2,\theta(\tau)}}{\sqrt{n}}, \hat{\theta}(\tau) - \frac{\xi_{\alpha/2,\theta(\tau)}}{\sqrt{n}}\right].$$

It is also possible to construct uniform confidence bands for $\tau \in (\delta, 1-\delta)$. In this case, we look for the smallest scalar c such that, under the bootstrap probability measure,

$$\Pr{}^*\left(\sup_{\tau \in (\delta, 1-\delta)} \left|\sqrt{n}(\hat{\theta}(\tau)^* - \hat{\theta}(\tau))\right| \leq c \,\middle|\, \{Y_i, D_i, D_{\delta,i}\}_{i=1}^n\right) \geq 1 - \alpha.$$

The unknown scalar c can be obtained by the simulation procedure outlined below.

Algorithm 2 (Bootstrap for θ)**.**

1. *Given the data* $\{Y_i, D_i, D_{\delta,i}\}_{i=1}^n$ *and a grid of values* $\{\tau_k\}_{k=1}^K \subset (\delta, 1-\delta)$*, compute* $\hat{\theta}(\tau_k)$ *as in* (3.21) *for each* $k = 1, \ldots, K$*.*

2. *Obtain B bootstrap samples of size n from* $\{Y_i, D_i, D_{\delta,i}\}_{i=1}^n$*, and compute*

$$\max_{k=1,\ldots,K} \left|\sqrt{n}(\hat{\theta}^b(\tau_k) - \hat{\theta}(\tau_k))\right|,$$

where $\hat{\theta}^b(\tau_k)$ *is computed as in* (3.21) *for* $b = 1, \ldots, B$ *and each* $k = 1, \ldots, K$*.*

3. *Obtain the* $(100\times(1-\alpha))\%$ *percentile of* $\left\{\max_{k=1,\dots,K}\left|\sqrt{n}(\hat{\theta}^b(\tau_k)-\hat{\theta}(\tau_k))\right|\right\}_{b=1}^{B}$. *This is denoted* $\xi_{1-\alpha,\theta}$.

The $1-\alpha$ confidence bands are then computed as

$$\mathscr{C}\mathscr{B}(\theta(\tau),\alpha)=\left[\hat{\theta}(\tau)-\frac{\xi_{1-\alpha,\theta}}{\sqrt{n}},\hat{\theta}(\tau)+\frac{\xi_{1-\alpha,\theta}}{\sqrt{n}}\right].$$

3.5.2 Bounds on the Global Effect

An important case is when we are interested in two conclusions $G_{\tau_1}>g$ and $G_{\tau_2}<g$ for $\tau_1\neq\tau_2$, both in $(\delta,1-\delta)$. This is the case in Figure 3.6. For the case of the global effect, by Theorem 3.2, the bounds are $\tau\mapsto F_a^{-1}(\tau-\delta U_{\tau_1})-F_Y^{-1}(\tau)$, and $\tau\mapsto F_a^{-1}(\tau-\delta L_{\tau_2})-F_Y^{-1}(\tau)$, for fixed values of U_{τ_1} and L_{τ_2}. The goal is to make inference on these bounds when U_{τ_1}, L_{τ_2}, F_a^{-1} and F_Y^{-1} are estimated.[10]

Define $B(U_{\tau_1};\tau):=F_a^{-1}(\tau-\delta U_{\tau_1})-F_Y^{-1}(\tau)$ and $B(L_{\tau_2};\tau):=F_a^{-1}(\tau-\delta L_{\tau_2})-F_Y^{-1}(\tau)$. The estimated counterparts are

$$\begin{pmatrix}\hat{B}(\hat{U}_{\tau_1};\tau)\\ \hat{B}(\hat{L}_{\tau_2};\tau)\end{pmatrix}=\begin{pmatrix}\hat{F}_a^{-1}(\tau-\delta\hat{U}_{\tau_1})-\hat{F}_Y^{-1}(\tau)\\ \hat{F}_a^{-1}(\tau-\delta\hat{L}_{\tau_2})-\hat{F}_Y^{-1}(\tau)\end{pmatrix},$$

where by (3.15), and (3.16) we have

$$\begin{pmatrix}\hat{U}_{\tau_1}\\ \hat{L}_{\tau_2}\end{pmatrix}=\begin{pmatrix}\min\{\max\{0,\hat{\theta}(\tau_1)\},1\}\\ \max\{\min\{0,\hat{\theta}(\tau_2)\},-1\}\end{pmatrix}.$$

To find the distributions of $\hat{U}_{\tau_1}$ and $\hat{L}_{\tau_2}$, we define the map $\phi:\ell^\infty(\delta,1-\delta)\mapsto[-1,0]\times$

[10]It is assumed that both U_{τ_1} and L_{τ_2} exist, in the sense that there is a robust region for the conclusions $G_{\tau_1}>g$ and $G_{\tau_2}<g$.

$[0,1]$ given by

$$\phi(H) = \begin{pmatrix} \min\{\max\{0, H(\tau_1)\}, 1\} \\ \max\{\min\{0, H(\tau_2)\}, -1\} \end{pmatrix}. \tag{3.22}$$

Though continuous, the composition of max and min (and vice versa) is not smooth. However, a form of differentiability, namely Hadamard directional differentiability, is still preserved. More importantly, the Delta method is still valid under this weaker differentiability notion. See Shapiro (1990), Dümbgen (1993), and, more recently and with applications to econometric theory, Fang and Santos (2019).

Theorem 3.5. *Under the Assumptions of Theorem 3.4,*

$$\sqrt{n}\begin{pmatrix} \hat{U}_{\tau_1} - U_{\tau_1} \\ \hat{L}_{\tau_2} - L_{\tau_2} \end{pmatrix} \rightsquigarrow \phi'_\theta(\mathbb{G}_\theta),$$

where

$$\phi'_\theta(\mathbb{G}_\theta) = \begin{pmatrix} \mathbb{G}_\theta(\tau_1)\mathbb{1}_{\{0<\theta(\tau_1)<1\}} + \max(0, \mathbb{G}_\theta(\tau_1))\mathbb{1}_{\{\theta(\tau_1)=0\}} + \min(0, \mathbb{G}_\theta(\tau_1))\mathbb{1}_{\{\theta(\tau_1)=1\}} \\ \mathbb{G}_\theta(\tau_2)\mathbb{1}_{\{-1<\theta(\tau_2)<0\}} + \min(0, \mathbb{G}_\theta(\tau_2))\mathbb{1}_{\{\theta(\tau_2)=0\}} + \max(0, \mathbb{G}_\theta(\tau_2))\mathbb{1}_{\{\theta(\tau_2)=-1\}} \end{pmatrix}. \tag{3.23}$$

It is important to point out that the distribution of $\phi'_\theta(\mathbb{G}_\theta)$ is *not* Gaussian. This is not only due to the presence of the min and max functions, but also because when $\theta(\tau_1) \notin [0,1]$ the first coordinate is degenerate in 0. The same comment applies to the second coordinate, which is degenerate when $\theta(\tau_2) \notin [-1,0]$. See Example 2.1 in Fang and Santos (2019) for a similar situation.

We need the following assumption in order to establish the Hadamard differentiability of the quantile process $\tau \mapsto F_a^{-1}(\tau)$ for $\tau \in (\delta, 1-\delta)$. Recall that the support of F_a is assumed to be $\mathscr{Y}$.

Assumption 3.6. *For some $\varepsilon > 0$, F_a is continuously differentiable in $[F_a^{-1}(\delta) - \varepsilon, F_a^{-1}(1-\delta) + \varepsilon] \subset \mathscr{Y}$ with strictly positive derivative f_a.*

When the bounds are viewed as a map in $\ell^\infty(\delta, 1-\delta) \times \ell^\infty(\delta, 1-\delta)$, we use "$\cdot$" to keep track of where the argument of the function should be, and we write

$$\sqrt{n}\begin{pmatrix} \hat{B}(\hat{U}_{\tau_1};\cdot) - B(U_{\tau_1};\cdot) \\ \hat{B}(\hat{L}_{\tau_2};\cdot) - B(L_{\tau_2};\cdot) \end{pmatrix} \tag{3.24}$$

Theorem 3.6. *Under Assumptions 3.3, 3.4, 3.5, and 3.6*

$$\sqrt{n}\begin{pmatrix} \hat{B}(\hat{U}_{\tau_1};\cdot) - B(U_{\tau_1};\cdot) \\ \hat{B}(\hat{L}_{\tau_2};\cdot) - B(L_{\tau_2};\cdot) \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{G}_{U_{\tau_1}} \\ \mathbb{G}_{L_{\tau_2}} \end{pmatrix}$$

a tight process in $\ell^\infty(\delta, 1-\delta) \times \ell^\infty(\delta, 1-\delta)$ given by

$$\begin{pmatrix} \mathbb{G}_{U_{\tau_1}} \\ \mathbb{G}_{L_{\tau_2}} \end{pmatrix} := \begin{pmatrix} -\frac{\mathbb{G}_a \circ F_a^{-1}(\,\cdot - \delta U_{\tau_1})}{f_a \circ F_a^{-1}(\,\cdot - \delta U_{\tau_1})} - \frac{\delta \phi_\theta'(\mathbb{G}_\theta)_2}{f_a \circ F_a^{-1}(\,\cdot - \delta U_{\tau_1})} - \frac{\mathbb{G}_Y \circ F_Y^{-1}(\cdot)}{f_Y \circ F_Y^{-1}(\cdot)} \\ -\frac{\mathbb{G}_a \circ F_a^{-1}(\,\cdot - \delta L_{\tau_2})}{f_a \circ F_a^{-1}(\,\cdot - \delta L_{\tau_2})} - \frac{\delta \phi_\theta'(\mathbb{G}_\theta)_1}{f_a \circ F_a^{-1}(\,\cdot - \delta L_{\tau_2})} - \frac{\mathbb{G}_Y \circ F_Y^{-1}(\cdot)}{f_Y \circ F_Y^{-1}(\cdot)} \end{pmatrix}, \tag{3.25}$$

where the map $\phi_\theta'(\mathbb{G}_\theta)$ is given in (3.23)*, and $\phi_\theta'(\mathbb{G}_\theta)_1$ and $\phi_\theta'(\mathbb{G}_\theta)_2$ are the first and second coordinates respectively.*

The limiting process in (3.25) is *not* Gaussian because of the presence of $\phi_\theta'(\mathbb{G}_\theta)$ given in Theorem 3.4. Hence, by Corollary 3.1 in Fang and Santos (2019), the standard bootstrap will fail. This means that if we attempt to construct confidence intervals in the usual way by resampling $\hat{B}(\hat{U}_{\tau_1};\tau)$ and $\hat{B}(\hat{L}_{\tau_2};\tau)$, we will not obtain correct asymptotic coverage. Instead, we use the numerical bootstrap of Hong and Li (2018, 2020). For a given τ, we write the map in

(3.24) as

$$\psi(F_a, F_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau) = \begin{pmatrix} F_a^{-1}(\tau - \delta \hat{U}_{\tau_1} - F_Y^{-1}(\tau) \\ F_a^{-1}(\tau - \delta \hat{L}_{\tau_2} - F_Y^{-1}(\tau) \end{pmatrix} \tag{3.26}$$

The idea is that we can approximate $(\mathbb{G}_{U_{\tau_1}}(\tau), \mathbb{G}_{L_{\tau_2}}(\tau))'$ using a standard bootstrap for $\hat{F}_a$ and $\hat{F}_Y$ and a numerical bootstrap for $\hat{U}_{\tau_1}$ and $\hat{L}_{\tau_2}$. First, fix $\hat{U}_{\tau_1}$ and $\hat{L}_{\tau_2}$, and let $\hat{F}_a^*$ and $\hat{F}_Y^*$ be the bootstrap counterparts of $\hat{F}_a$ and $\hat{F}_Y$. Then, define

$$\hat{\psi}^{\bullet}_{aY}(\hat{F}_a^*, \hat{F}_Y^*; \tau) = \sqrt{n}\left(\psi(\hat{F}_a^*, \hat{F}_Y^*, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau) - \psi(\hat{F}_a, \hat{F}_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau)\right). \tag{3.27}$$

Now, fix $\hat{F}_a$ and $\hat{F}_Y$, let $\theta(\tau_1)^*$ and $\hat{\theta}(\tau_2)^*$ be the bootstrap counterparts of $\hat{\theta}(\tau_1)$ and $\hat{\theta}(\tau_2)$, and define the perturbed parameters

$$\hat{\theta}(\tau_1)^p := \hat{\theta}(\tau_1) + \varepsilon_n \sqrt{n}(\hat{\theta}(\tau_1)^* - \hat{\theta}(\tau_1)), \tag{3.28}$$

and

$$\hat{\theta}(\tau_2)^p := \hat{\theta}(\tau_2) + \varepsilon_n \sqrt{n}(\hat{\theta}(\tau_2)^* - \hat{\theta}(\tau_2)), \tag{3.29}$$

where the sequence ε_n is constrained to satisfy $\varepsilon_n \to 0$ and $\varepsilon_n \sqrt{n} \to \infty$, as $n \to \infty$. In the empirical application we set $\varepsilon_n = n^{-1/3}$. Define, the perturbed version of U_{τ_1} and L_{τ_2} as

$$\begin{pmatrix} \hat{U}^p_{\tau_1} \\ \hat{L}^p_{\tau_2} \end{pmatrix} = \begin{pmatrix} \min\{\max\{0, \hat{\theta}(\tau_1)^p\}, 1\} \\ \max\{\min\{0, \hat{\theta}(\tau_2)^p\}, -1\} \end{pmatrix}. \tag{3.30}$$

Then, define

$$\hat{\psi}^{\bullet}_{UL}(\hat{U}^p_{\tau_1}, \hat{L}^p_{\tau_2}; \tau) = \frac{1}{\varepsilon_n}\left(\psi(\hat{F}_a, \hat{F}_Y, \hat{U}^p_{\tau_1}, \hat{L}^p_{\tau_2}; \tau) - \psi(\hat{F}_a, \hat{F}_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau)\right). \tag{3.31}$$

The approximation to $(\mathbb{G}_{U_{\tau_1}}(\tau), \mathbb{G}_{L_{\tau_2}}(\tau))'$ is given by the distribution of

$$\hat{\psi}^{\bullet}_{aY}(\hat{F}^*_a, \hat{F}^*_Y; \tau) + \hat{\psi}^{\bullet}_{UL}(\hat{U}^p_{\tau_1}, \hat{L}^p_{\tau_2}; \tau),$$

which, in turn, is approximated by the simulated procedure below.

Algorithm 3 (Bootstrap for $\hat{B}(\hat{U}_{\tau_1}; \tau)$ and $\hat{B}(\hat{L}_{\tau_2}; \tau)$)**.**

1. *Given the data* $\left\{Y_i, D_i, D_{\delta,i}\right\}_{i=1}^n$*, compute* $\psi(\hat{F}_a, \hat{F}_Y, \hat{\theta}(\tau_1), \hat{\theta}(\tau_2); \tau)$ *given in* (3.26).

2. *Obtain B bootstrap samples of size n from* $\left\{Y_i, D_i, D_{\delta,i}\right\}_{i=1}^n$.

3. *For* $b = 1, \ldots, B$*, following* (3.27)*, compute*

$$\hat{\psi}^{\bullet}_{aY}(\hat{F}^b_a, \hat{F}^b_Y; \tau) = \sqrt{n}\left(\psi(\hat{F}^b_a, \hat{F}^b_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau) - \psi(\hat{F}_a, \hat{F}_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau)\right). \tag{3.32}$$

4. *For* $b = 1, \ldots, B$*, following* (3.21)*, compute* $\hat{\theta}(\tau_1)^b$ *and* $\hat{\theta}(\tau_2)^b$*. Following* (3.28) *and* (3.29)*, compute the perturbed parameters as*

$$\hat{\theta}(\tau_1)^{p,b} := \hat{\theta}(\tau_1) + \varepsilon_n\sqrt{n}(\hat{\theta}(\tau_1)^b - \hat{\theta}(\tau_1)),$$

and

$$\hat{\theta}(\tau_2)^{p,b} := \hat{\theta}(\tau_2) + \varepsilon_n\sqrt{n}(\hat{\theta}(\tau_2)^b - \hat{\theta}(\tau_2))$$

Following (3.30) *compute*

$$\begin{pmatrix} \hat{L}^{p,b}_{\tau_2} \\ \hat{U}^{p,b}_{\tau_1} \end{pmatrix} = \begin{pmatrix} \max\{\min\{0, \hat{\theta}(\tau_2)^{p,b}\}, -1\} \\ \min\{\max\{0, \hat{\theta}(\tau_1)^{p,b}\}, 1\} \end{pmatrix}.$$

5. *For* $b = 1, \ldots, B$, *following* (3.31), *compute*

$$\hat{\psi}^{\bullet}_{UL}(\hat{U}^{p,b}_{\tau_1}, \hat{L}^{p,b}_{\tau_2}; \tau) = \frac{1}{\varepsilon_n}\left(\psi(\hat{F}_a, \hat{F}_Y, \hat{U}^{p,b}_{\tau_1}, \hat{L}^{p,b}_{\tau_2}; \tau) - \psi(\hat{F}_a, \hat{F}_Y, \hat{U}_{\tau_1}, \hat{L}_{\tau_2}; \tau)\right).$$

6. *For* $b = 1, \ldots, B$, *define*

$$\hat{\psi}'(b, \tau) = \hat{\psi}^{\bullet}_{aY}(\hat{F}^b_a, \hat{F}^b_Y; \tau) + \hat{\psi}^{\bullet}_{UL}(\hat{U}^{p,b}_{\tau_1}, \hat{L}^{p,b}_{\tau_2}; \tau). \tag{3.33}$$

7. *Obtain the* $(100 \times \alpha/2)\%$ *and* $(100 \times (1-\alpha/2))\%$ *percentiles from the first coordinate of* (3.33). *These are denoted* $\xi_{\alpha/2,\tau,U_{\tau_1}}$ *and* $\xi_{1-\alpha/2,\tau,U_{\tau_1}}$.

8. *Obtain the* $(100 \times \alpha/2)\%$ *and* $(100 \times (1-\alpha/2))\%$ *percentiles from the second coordinate of* (3.33). *These are denoted* $\xi_{\alpha/2,\tau,L_{\tau_2}}$ *and* $\xi_{1-\alpha/2,\tau,L_{\tau_2}}$.

The $1-\alpha$ confidence intervals are then computed as

$$\mathscr{CI}(B(U_{\tau_1}; \tau), \alpha) = \left[\hat{B}(\hat{U}_{\tau_1}; \tau) - \frac{\xi_{1-\alpha/2,\tau,U_{\tau_1}}}{\sqrt{n}}, \hat{B}(\hat{U}_{\tau_1}; \tau) - \frac{\xi_{\alpha/2,\tau,U_{\tau_1}}}{\sqrt{n}}\right],$$
$$\mathscr{CI}(B(L_{\tau_2}; \tau), \alpha) = \left[\hat{B}(\hat{L}_{\tau_2}; \tau) - \frac{\xi_{1-\alpha/2,\tau,L_{\tau_2}}}{\sqrt{n}}, \hat{B}(\hat{L}_{\tau_2}; \tau) - \frac{\xi_{\alpha/2,\tau,L_{\tau_2}}}{\sqrt{n}}\right].$$

The simultaneous $1-\alpha$ confidence intervals, by the Bonferroni correction,[11] are given by the Cartesian product

$$\mathscr{CI}(B(U_{\tau_1}; \tau), B(L_{\tau_2}; \tau), \alpha) = \mathscr{CI}(B(U_{\tau_1}; \tau), \alpha/2) \times \mathscr{CI}(B(L_{\tau_2}; \tau), \alpha/2).$$

Alternatively, simultaneous $1-\alpha$ confidence intervals can be constructed using a lower confidence interval for $B(U_{\tau_1}; \tau)$: $\hat{B}(\hat{U}_{\tau_1}; \tau) - \frac{\xi_{1-\alpha/2,\tau,U_{\tau_1}}}{\sqrt{n}}$, and upper confidence interval for $B(L_{\tau_2}; \tau)$: $\hat{B}(\hat{L}_{\tau_2}; \tau) - \frac{\xi_{\alpha/2,\tau,L_{\tau_2}}}{\sqrt{n}}$.

[11] If we want simultaneous $1-\alpha$ confidence intervals, for each coordinate the confidence intervals must be constructed at the $1-\alpha/2$ level.

We can construct uniform confidence bands for $\tau \in (\delta, 1-\delta)$ in the following way.

Algorithm 4 (Bootstrap for $\hat{B}(\hat{U}_{\tau_1};\cdot)$ and $\hat{B}(\hat{L}_{\tau_2};\cdot)$)**.**

1. *Given a grid of values* $\{\tau_k\}_{k=1}^K \subset (\delta, 1-\delta)$*, following* (3.33)*, compute for* $b = 1, \ldots, B$

$$\hat{\psi}'(b) = \max_{k=1,\ldots,K} \left| \hat{\psi}^{\bullet}_{aY}(\hat{F}^b_a, \hat{F}^b_Y; \tau_k) + \hat{\psi}^{\bullet}_{UL}(\hat{U}^{p,b}_{\tau_1}, \hat{L}^{p,b}_{\tau_2}; \tau_k) \right| \tag{3.34}$$

2. *Obtain the* $(100 \times (1-\alpha))\%$ *percentile from the first coordinate of* (3.34)*. This is denoted* $\xi_{1-\alpha, U_{\tau_1}}$.

3. *Obtain the* $(100 \times (1-\alpha))\%$ *percentile from the second coordinate of* (3.34)*. This is denoted* $\xi_{1-\alpha, L_{\tau_2}}$.

The one-sided or two-sided $1-\alpha$ confidence bands are computed as before.

3.5.3 Quantile Breakdown Frontier: Marginal Effect

The quantile breakdown frontier for the sign of the marginal effect is given by (see (3.20)) the map $\tau \mapsto F_{Y|D=0}(F_Y^{-1}(\tau)) - F_{Y|D=1}(F_Y^{-1}(\tau))$, and the estimated counterpart is $\hat{\theta}(\tau) = \hat{F}_{Y|D=0}(\hat{F}_Y^{-1}(\tau)) - \hat{F}_{Y|D=1}(\hat{F}_Y^{-1}(\tau))$, where

$$\hat{F}_{Y|D=0}(y) := \frac{\sum_{i=1}^n \mathbb{1}\{Y_i \leq y\}(1-D_i)}{\sum_{i=1}^n (1-D_i)},$$

and

$$\hat{F}_{Y|D=1}(y) := \frac{\sum_{i=1}^n \mathbb{1}\{Y_i \leq y\} D_i}{\sum_{i=1}^n D_i}.$$

As before, we want to investigate the weak convergence of $\sqrt{n}(\hat{\theta} - \theta)$ in $\ell^\infty(0,1)$:

$$\sqrt{n}(\hat{\theta} - \theta) = \sqrt{n}\left(\hat{F}_{Y|D=0} \circ \hat{F}_Y^{-1} - \hat{F}_{Y|D=1} \circ \hat{F}_Y^{-1} - \left(F_{Y|D=0} \circ F_Y^{-1} - F_{Y|D=1} \circ F_Y^{-1}\right)\right).$$

Recall that the bounds on the marginal effect can be computed for any $\tau \in (0,1)$, as opposed to the global effect, where we are constrained to $\tau \in (\delta, 1-\delta)$. The main assumption is

Assumption 3.7 (Functional CLT)**.** *The following multivariate functional central limit theorem holds*

$$\sqrt{n}\begin{pmatrix} \hat{F}_Y - F_Y \\ \hat{F}_{Y|D=0} - F_{Y|D=0} \\ \hat{F}_{Y|D=1} - F_{Y|D=1} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{G}_Y \\ \mathbb{G}_0 \\ \mathbb{G}_1 \end{pmatrix},$$

where $\mathbb{G}_Y$, $\mathbb{G}_0$, and $\mathbb{G}_1$ are Brownian bridges in $\ell^\infty(\mathscr{Y})$, where $\mathscr{Y}$ is the common support of Y, $Y|D=0$, and $Y|D=1$.

The next assumption is needed to establish the Hadamard differentiability of the composition map, and the quantile process.

Assumption 3.8 (Conditions for Hadamard Differentiability)**.**

1. *The distribution functions $F_{Y|D=0}(y)$ and $F_{Y|D=1}(y)$ are differentiable, with uniformly continuous and bounded derivatives on their support $\mathscr{Y}$. The derivatives are $f_{Y|D=0}(y)$ and $f_{Y|D=1}(y)$ respectively.*
2. *The support $\mathscr{Y}$ is the compact set $[y_l, y_u]$.*
3. *$F_Y(y)$ is continuously differentiable on $\mathscr{Y}$ with strictly positive derivative f_Y.*

Theorem 3.7 (Asymptotic Distribution of QBF for Marginal Effect)**.** *Under Assumptions 3.7 and 3.8*

$$\begin{aligned}\sqrt{n}(\hat{\theta} - \theta) &= \sqrt{n}\left(\hat{F}_{Y|D=0} \circ \hat{F}_Y^{-1} - \hat{F}_{Y|D=1} \circ \hat{F}_Y^{-1} - \left(F_{Y|D=0} \circ F_Y^{-1} - F_{Y|D=1} \circ F_Y^{-1}\right)\right) \\ &\rightsquigarrow \mathbb{G}_{0,Y} - \mathbb{G}_{1,Y},\end{aligned}$$

where, for $d = 0,1$, $\mathbb{G}_{d,Y} := \mathbb{G}_d \circ F_Y^{-1} - f_{Y|D=d} \circ F_Y^{-1} \cdot \frac{\mathbb{G}_Y \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}}$ *are tight Gaussian elements of* $\ell^\infty(0,1)$.

Confidence intervals/bands can be constructed following the same procedures outlined in Algorithms 1 and 2, because by Theorem 23.9 in van der Vaart (1998), the empirical bootstrap is valid. We skip the details to avoid repetition.

3.6 Empirical application: What do unions do?

There is an extensive literature that studies unions and inequality. A recent contribution by Farber et al. (2020) contains a review of the literature. In our empirical application, in particular, we look at how unions affect the distribution of wages for *all* workers. Unions can have a variety of effects on the distribution of wages. As argued by Freeman (1980), unions can raise the wages of unionized workers relative to non-unionized workers, possibly through more bargaining power. So, if higher paid workers unionize, the dispersion of wages can increase, but if lower paid workers unionize, the dispersion of wages can decrease. Furthermore, within a given industry, the union can reduce the dispersion of wages by standardizing the wages. This will impact the distribution of wages more or less depending on the size of the industry and the wages it pays.

A key difficulty in identifying the causal effect of unions on wages is that selection into unions is non-random. Hence, any measurement of the union premium, the difference in wages between similar union and nonunion workers, will be biased for the causal effect. Indeed, this has been a long standing concern of labor economists. With respect to selection into unions, Card (1996) argues that unionized workers with low observed skills, tend to have high unobserved skills. The reverse happens with high skilled unionized workers: they tend to have low unobservable skills. Due to this selection bias, it might be impossible for a policy maker to device a policy where the *newly* unionized workers are selected in a way such that they are drawn from the distribution of the *already* unionized workers.

Using the techniques developed in this chapter, we are going to consider the effect of both globally and marginally expanding union coverage. We will explicitly allow for non-random selection into unions. Moreover, as opposed to Firpo et al. (2009), we will not assume distributional invariance: the distribution of the *newly* unionized workers can be different from the distribution of the *already* unionized workers. That is, we do not use any imputation method to impute the union premium of the newly unionized workers

Following Freeman (1980), Card (2001) and Card et al. (2004) we consider a two sector economy. Each worker has a well-defined pair of potential (log) wages: $Y_i(1)$ for the unionized sector and $Y_i(0)$ for the nonunionized sector. Under Assumption 3.1, and for any policy D_δ, we have the following classification of individuals:

Table 3.2. Clasification of individuals

	$D_\delta = 0$	$D_\delta = 1$
$D = 0$	*nonunionized*	*newly unionized*
$D = 1$	-	*unionized*

The relevant unobserved distribution is then $F_{Y(1)|newly\ unionized}$: the union wages of the newly unionized workers. So, we look at departures of $F_{Y(1)|newly\ unionized}$ from $F_{Y(1)|unionized}$, which is observed. This difference is what we refer to as the policy selection bias.

Using the data in Firpo et al. (2009) we estimate the quantile breakdown frontier for marginal and global effects of different type of policies on the distribution of real log hourly wages. We use the 1983-1985 Outgoing Rotation Group (ORG) Supplement of the Current Population Survey. Our sample consists of 266,956 observations on U.S. males. See Lemieux (2006) for more details about the data.

The unionization rate in the dataset is 0.26. Figure 3.10 shows the typical hump-shaped pattern of the unionization rates by quantiles of the distribution of wages. For lower quantiles, unionization rates are quite low. They peak in the past the middle of the distribution and then drop at the higher quantiles. We will analyze a randomized policy and a non-randomized policy.

In the first case, we will analyze the policy that marginally increases unionization by selecting workers at random. We will look at the quantile breakdown frontier for the sign of the marginal effect. That is, we set $g = 0$ and look at whether the marginal effect is positive or negative. Figure 3.11 shows the result for a grid of $\tau \in (0.1, 0.9)$, along with 95% pointwise confidence intervals and uniform confidence bands. We can see that along almost all quantiles, the quantile breakdown frontier is positive, and it peak at around 0.27 for $\tau = 0.4$. This means that if the selection bias due to U is greater than 0.27, then the conclusion $M_\tau > 0$ does not hold for any τ.

In the second case, we will analyze a non-randomized policy. Consider a 10% increase in the unionization rate by unionizing workers whose wages are below the $.10/(1-p)$-quantile ≈ 0.14-quantile of the wages of the nonunionized sector. In the notation of this chapter, we have $D = 1$ if a worker is unionized, $D_\delta = 1$ if a worker is unionized under the policy, Y is (log) wage, and $\delta = 0.1$. That is, D_δ is given by

$$D_\delta = \begin{cases} 1 & \text{if } D = 1 \\ 1 & \text{if } D = 0 \text{ and } Y \leq F^{-1}_{Y|D=0}(0.14) \\ 0 & \text{otherwise} \end{cases}$$

This guarantees that the unionization rate increases by roughly 10%. Indeed, the mean of D_δ is now 0.36. Figure 3.12 shows the quantile breakdown frontiers for $g_\tau = 0.1$ for a grid of $\tau \in (0.1, 0.9)$. This is the empirical counterpart of the right side of Figure 3.5. Pointwise confidence intervals (shaded) and uniform confidence bands (dashed) are also shown, both at the 95% level. Since the dependent variable is log wages, $g_\tau = 0.1$ amounts to a 10% change in wages for a given quantile.

For lower quantiles, if we want the policy to result in an increase of wages *higher* than 10%, then the departure from point identification is given by U_τ in the positive part of the curve: for example, for the 20$^{\text{th}}$ quantile, $U_{.2} \approx 0.45$. For higher quantiles, if we want the policy to result in changes of wages *lower* than 10%, then the maximum departure from point identification, L_τ, is given by the negative part of the curve. For example, for the 80th quantile, $L_{.8} \approx -0.36$. In terms of our notation, if we are interested in the conclusions $G_{.2,D_{.1}} > 0.1$ and $G_{.8,D_{.1}} < 0.1$, then the robust region is

$$\{(L,U) : -1 \leq L \leq -0.36 \text{ and } 0 \leq U \leq 0.45\}$$

Recall that U and L come from Assumption 3.2:

$$L \leq F_{Y(1)|newly\ unionized}(y) - F_{Y(1)|unionized}(y) \leq U.$$

So if we are interested in the 20th and 80th quantile, we need

$$-0.36 \leq F_{Y(1)|newly\ unionized}(y) - F_{Y(1)|unionized}(y) \leq 0.45. \tag{3.35}$$

for the conclusions to hold. This does not rule out either direction of first-order stochastic domination, but it does put a bound on it. Since $F_{Y(1)|unionized}(y)$ can be estimated, then simulation exercises can be carried out on possible CDFs that satisfy (3.35), *i.e.*, they are not too far away from the empirical counterpart of $F_{Y(1)|unionized}(y)$. Figure 3.13 shows the estimated bounds for the global effect when setting $L_{.8} \approx -0.36$ and $U_{.2} \approx 0.45$. For $\tau = 0.2$, we can see that the identified region lies above 0.10, and for $\tau = 0.8$, the identified region lies below 0.10. Pointwise confidence intervals (shaded) and uniform confidence bands (dashed) are also shown, both at the 95% level.

We repeat the same exercise for the global effect, this time for $g = 0.05$. We keep $\delta = 0.1$. The quantile breakdown frontier and the bounds on the global effect can be seen in Figures 3.14 and 3.15. At the 20th quantile, $U_{.2} \approx 0.67$, while at the 80th quantile, $L_{.8} \approx -0.05$. This means that the hypothesis $G_{.8,D_{.1}} < 0.05$ is not very robust: any policy selection bias above given L in $[-1, -0.05)$ result in identification regions for $G_{.8,D_{.1}}$ that contain values greater than 0.05.

Figures 3.13 and 3.15 show that, because of the continuity of the quantile breakdown frontier, when we focus on conclusions at the 20th and 80th quantiles, we are also deriving bounds for the global effect at other quantiles. Thus, in Figure 3.13, we can see that the global effect, which is consistent with $G_{.2,D_{.1}} > 0.1$ and $G_{.8,D_{.1}} < 0.1$ is positive up to $\tau = 0.6$. In other words, the combinations of L and U that ensure that $G_{.2,D_{.1}} > 0.1$ and $G_{.8,D_{.1}} < 0.1$, imply that $G_{\tau,D_{.1}} > 0$ for $\tau \in (0.1, 0.6)$.

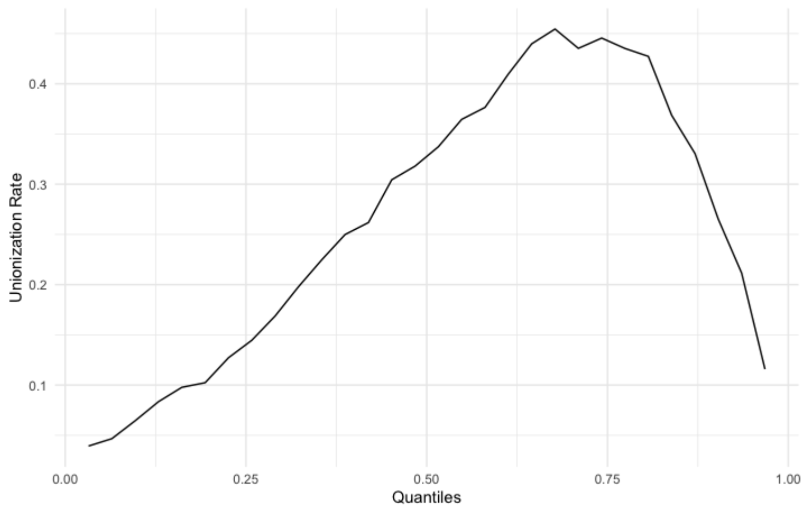

Figure 3.10. Unionization rates by quantiles of the distribution of wages.

3.7 Conclusion

In this paper we show how to perform a sensitivity analysis on the effect of counterfactual policies on the quantiles of an outcome of interest. We focus on counterfactual policies which increase the proportion of treated individuals and obtain partial identified sets for both global and marginal effects on the unconditional quantiles. In the former, the increase δ in the proportion is fixed, while in the latter goes to 0. By dropping the standard distributional invariance assumption, we are able to broaden the scope of policies that can be analyzed. Our partial identification results are used to perform a sensitivity analysis based on the departure from point identification. The sensitivity analysis is greatly simplified by the introduction of the quantile breakdown frontier, a curve that quantifies the maximum amount of selection bias compatible with a given conclusion

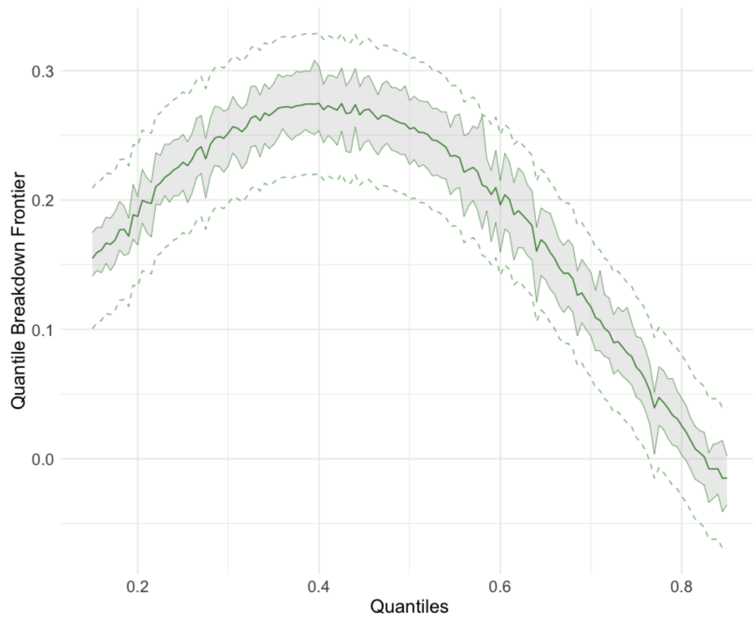

Figure 3.11. Quantile Breakdown Frontier for the sign of the marginal effect.

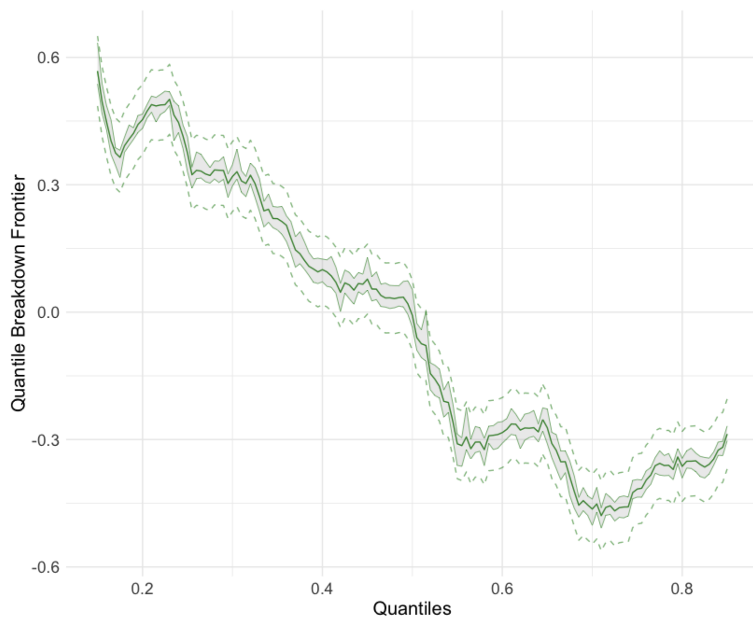

Figure 3.12. Quantile Breakdown Frontier for the global effect and $g = 0.1$.

at each quantile. A further use of the quantile breakdown frontier, is to bound the global effect curve in order for it to be consistent with a set of desired conclusions.

Our empirical application takes another look at the relationship between unions and

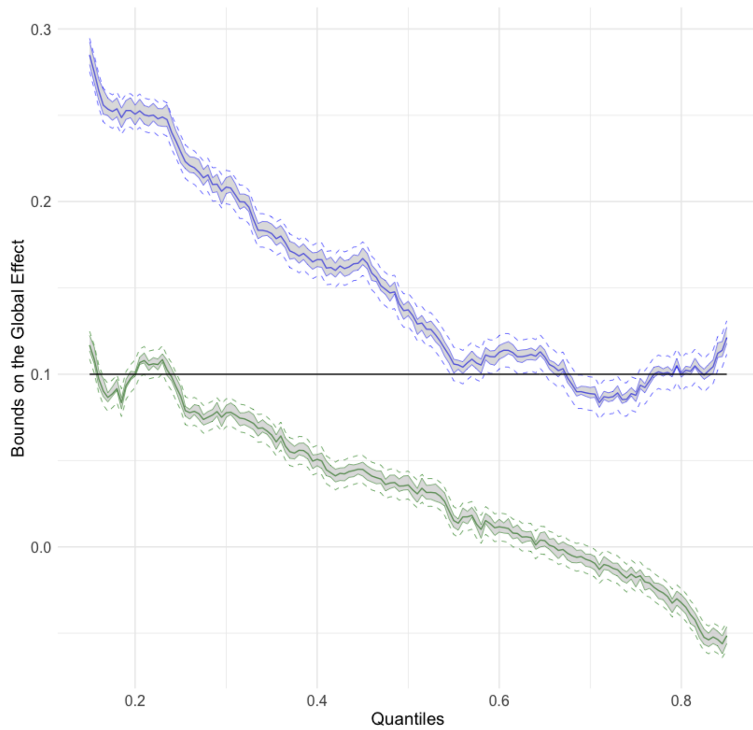

Figure 3.13. Bounds on the global effect for $L_{.8} \approx -0.36$ and $U_{.2} \approx 0.45$ and $g = 0.1$.

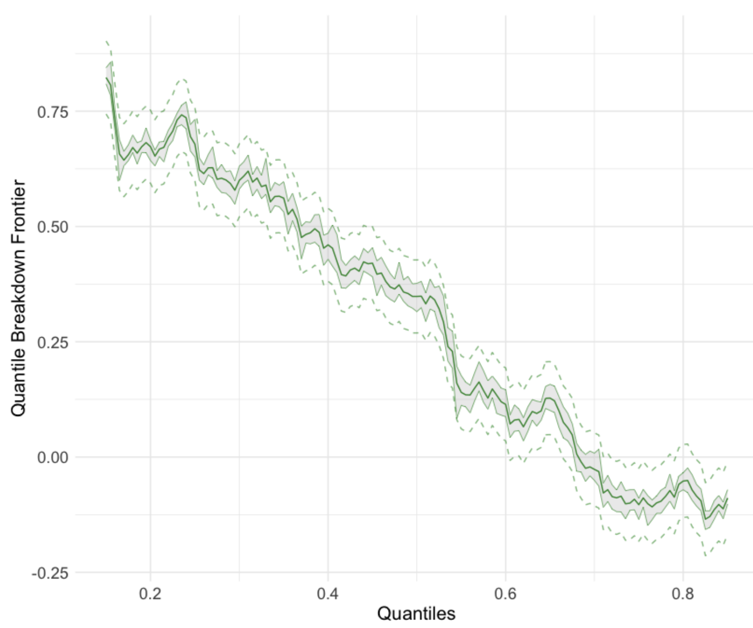

Figure 3.14. Quantile Breakdown Frontier for the global effect and $g = 0.05$.

inequality. In particular, we perform a sensitivity analysis on a policy that increases unionization by 10%. This is done by selecting nonunionized workers who are below a certain threshold of

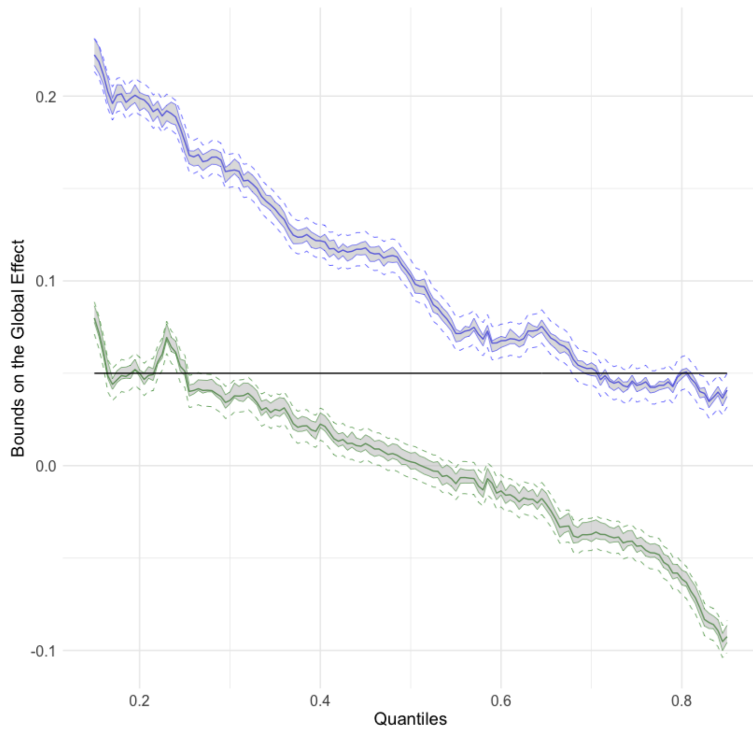

Figure 3.15. Bounds on the global effect for $L_{.8} \approx -0.05$ and $U_{.2} \approx 0.67$ and $g = 0.05$.

income. We then look at the effect of this policy on the 20th and 80th quantiles of the distribution of wages. We are interested in the following conclusion: the change in the 20th quantile of wages is greater than 10%, while the change at the 80th quantile is less than 10%. We derive the values of selection bias consistent with the conclusion. Our results show that this policy is consistent with moderate values of selection bias.

Chapter 3 is currently being prepared for submission for publication of the material. The dissertation author, Julián Martínez-Iriarte, was the sole author of this material.

Appendix A

Supplementary Proofs

Proof of Theorem 1.1. Let $\Gamma_\tau[F]$ be the τ-quantile of F. The Hadamard derivative at F is (See Lemma 21.3 in van der Vaart (1998))

$$\Gamma'_{\tau,F}[h] = -\frac{r(F^{-1}(\tau))}{f(F^{-1}(\tau))}.$$

for any $h \in D[-\infty,\infty]$ continuous at $F^{-1}(\tau)$.[1] We write the marginal effect as

$$\begin{aligned}
M_{\tau,\mathscr{D}} &= \lim_{\delta\downarrow 0} \frac{\Gamma_\tau\left[F_{Y_{D_\delta}}\right] - \Gamma_\tau[F_Y]}{\delta} \\
&= \lim_{\delta\downarrow 0} \frac{\Gamma_\tau\left[F_{Y_{D_0}} + \delta\left(\frac{F_{Y_{D_\delta}} - F_{Y_{D_0}}}{\delta}\right)\right] - \Gamma_\tau[F_Y]}{\delta} \\
&= \lim_{\delta\downarrow 0} \frac{\Gamma_\tau\left[F_Y + \delta\left(\frac{F_{Y_{D_\delta}} - F_Y}{\delta}\right)\right] - \Gamma_\tau[F_Y]}{\delta} \\
&= \Gamma'_{\tau,F_Y}[\dot{F}_{Y,\mathscr{D}}] \\
&= \frac{\dot{F}_{Y,\mathscr{D}}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))}.
\end{aligned}$$

[1] For $[a,b] \subset [-\infty,\infty]$, $D[a,b]$ is the Skorohod space: the set of all real-valued cadlag functions: right continuous with left limits everywhere in $[a,b]$. $D[a,b]$ is equipped with the uniform norm: $\|x\|_\infty := \sup_{t\in[a,b]} |x(t)|$.

The third equality follows from $F_{Y_{D_0}} = F_Y$. The fourth equality follows from

$$\limsup_{\delta \downarrow 0} \sup_{y \in \mathscr{Y}} \left| \frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} - \dot{F}_{Y,\mathscr{D}}(y) \right| = 0,$$

which is required by Lemma 21.3 in van der Vaart (1998).

□

Proof of Lemma 1.1. Using our notation, we follow Kaplan (2020) closely. We start with the model given in (1.1).

$$\begin{aligned}
F_{Y_{D_\delta}}(y) &= (p+\delta)\int_{\mathscr{X}_1} F_{Y_{D_\delta}|D_\delta=1,X=x}(y)dF_{X|D_\delta=1}(x) \\
&\quad + (1-p-\delta)\int_{\mathscr{X}_0} F_{Y_{D_\delta}|D_\delta=0,X=x}(y)dF_{X|D_\delta=0}(x) \\
&= (p+\delta)\int_{\mathscr{X}_1}\int_{\mathscr{Y}} 1\{r(1,x,U)\le y\}\, dF_{U|D_\delta=1,X=x}(u)dF_{X|D_\delta=1}(x) \\
&\quad + (1-p-\delta)\int_{\mathscr{X}_0}\int_{\mathscr{Y}} 1\{r(0,x,U)\le y\}\, dF_{U|D_\delta=0,X=x}(u)dF_{X|D_\delta=0}(x) \\
&= (p+\delta)\int_{\mathscr{X}_1}\int_{\mathscr{Y}} 1\{r(1,x,U)\le y\}\, dF_{U|D=1,X=x}(u)dF_{X|D_\delta=1}(x) \\
&\quad + (1-p-\delta)\int_{\mathscr{X}_0}\int_{\mathscr{Y}} 1\{r(0,x,U)\le y\}\, dF_{U|D=0,X=x}(u)dF_{X|D_\delta=0}(x) \\
&= (p+\delta)\int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y)dF_{X|D_\delta=1}(x) + (1-p-\delta)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)dF_{X|D_\delta=0}(x).
\end{aligned}$$

□

Proof of Theorem 1.2. We need to find the limit of

$$\begin{aligned}
\frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} &= p\int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y)d\left(\frac{F_{X|D_\delta=1}(x) - F_{X|D=1}(x)}{\delta}\right) \\
&\quad + (1-p)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)d\left(\frac{F_{X|D_\delta=0}(x) - F_{X|D=0}(x)}{\delta}\right) \\
&\quad + \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y)dF_{X|D_\delta=1}(x) - \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)dF_{X|D_\delta=0}(x),
\end{aligned}$$

and show that convergence holds uniformly in $y \in \mathscr{Y}$.

We do this term by term. Consider the first term

$$p \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) d \left(\frac{F_{X|D_\delta=1}(x) - F_{X|D=1}(x)}{\delta} \right)$$

Under Assumption 1.3(2) we can write this as

$$p \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) \left(\frac{f_{X|D_\delta=1}(x) - f_{X|D=1}(x)}{\delta} \right) dx \tag{A.1}$$

Now,

$$f_{X|D_\delta=1}(x) = \Pr(D_\delta = 1|X = x) \frac{f_X(x)}{p+\delta}$$

and

$$f_{X|D=1}(x) = \Pr(D = 1|X = x) \frac{f_X(x)}{p}$$

So we have

$$\begin{aligned}
f_{X|D_\delta=1}(x) - f_{X|D=1}(x) &= \Pr(D_\delta = 1|X = x) \frac{f_X(x)}{p+\delta} - \Pr(D = 1|X = x) \frac{f_X(x)}{p} \\
&= \frac{f_X(x)}{p+\delta} \left(\Pr(D_\delta = 1|X = x) - \Pr(D = 1|X = x) \right) \\
&+ \Pr(D = 1|X = x) \left(\frac{f_X(x)}{p+\delta} - \frac{f_X(x)}{p} \right) \\
&= \frac{f_X(x)}{p+\delta} \left(\Pr(D_\delta = 1|X = x) - \Pr(D = 1|X = x) \right) \\
&- \Pr(D = 1|X = x) f_X(x) \frac{\delta}{(p+\delta)p}
\end{aligned}$$

Dividing through by δ, we obtain

$$\frac{f_{X|D_\delta=1}(x)-f_{X|D=1}(x)}{\delta}=\frac{f_X(x)}{p+\delta}\frac{\Pr(D_\delta=1|X=x)-\Pr(D=1|X=x)}{\delta}$$
$$-\Pr(D=1|X=x)\frac{f_X(x)}{(p+\delta)p}$$

Plugging this back in (A.1), we obtain

$$\begin{aligned}&p\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)d\left(\frac{F_{X|D_\delta=1}(x)-F_{X|D=1}(x)}{\delta}\right)\\&=\frac{p}{p+\delta}\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)\frac{\Pr(D_\delta=1|X=x)-\Pr(D=1|X=x)}{\delta}dF_X(x)\\&-\frac{1}{p+\delta}\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)\Pr(D=1|X=x)dF_X(x)\end{aligned}$$

Under Assumption 1.3(4), we can use dominated convergence theorem to pass the limit, and conclude (pointwise in y) that

$$\begin{aligned}&\lim_{\delta\to 0}p\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)d\left(\frac{F_{X|D_\delta=1}(x)-F_{X|D=1}(x)}{\delta}\right)\\&=\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)\dot{P}(x)dF_X(x)-F_{Y|D=1}(y)\end{aligned} \tag{A.2}$$

To show that convergence is uniform in $y\in\mathscr{Y}$, we show that each term in (A.2) separately converges uniformly. To alleviate notation, recall that $P(x):=\Pr(D=1|X=x)$ and that $P_\delta(x):=\Pr(D_\delta=1|X=x)$. So, for the first term (ignoring the $\frac{p}{p+\delta}$) factor, we have

$$\left|\int_{\mathscr{X}_1}F_{Y|D=1,X=x}(y)\left[\frac{P_\delta(x)-P(x)}{\delta}-\dot{P}(x)\right]dF_X(x)\right|\leq\int_{\mathscr{X}_1}\left|\frac{P_\delta(x)-P(x)}{\delta}-\dot{P}(x)\right|dF_X(x)$$

which does not depend on y. Thus, convergence holds uniformly. For the second term, the dependence on δ is given by the leading factor $\frac{1}{p+\delta}$. Thus, uniform convergence holds trivially, since $F_{Y|D=1}(y)$ is bounded.[2]

[2]To see this, consider a real valued sequence $a_n\to a$, and a real-valued function $g(y)$. Then $|a_ng(y)-ag(y)|\leq$

Now consider the second term

$$(1-p)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)d\left(\frac{F_{X|D_\delta=0}(x)-F_{X|D=0}(x)}{\delta}\right)$$

which we can write as

$$(1-p)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)\left(\frac{f_{X|D_\delta=0}(x)-f_{X|D=0}(x)}{\delta}\right)dx \tag{A.3}$$

because of Assumption 1.3(2). Now

$$\begin{aligned} f_{X|D_\delta=0}(x) &= \Pr(D_\delta=0|X=x)\frac{f_X(x)}{1-p-\delta} \\ &= \frac{f_X(x)}{1-p-\delta} - \Pr(D_\delta=1|X=x)\frac{f_X(x)}{1-p-\delta} \end{aligned}$$

and

$$\begin{aligned} f_{X|D=0}(x) &= \Pr(D=0|X=x)\frac{f_X(x)}{1-p} \\ &= \frac{f_X(x)}{1-p} - \Pr(D=1|X=x)\frac{f_X(x)}{1-p} \end{aligned}$$

$|a_n-a|\sup|g(y)|$. So we need $g(y)$ to be bounded.

Taking the difference, we get

$$
\begin{aligned}
f_{X|D_\delta=0}(x) - f_{X|D=0}(x) &= \frac{f_X(x)}{1-p-\delta} - \frac{f_X(x)}{1-p} \\
&\quad + \Pr(D=1|X=x)\frac{f_X(x)}{1-p} - \Pr(D_\delta=1|X=x)\frac{f_X(x)}{1-p-\delta} \\
&= f_X(x)\frac{\delta}{(1-p-\delta)(1-p)} \\
&\quad - \frac{f_X(x)}{(1-p-\delta)}\left[\Pr(D_\delta=1|X=x) - \Pr(D=1|X=x)\right] \\
&\quad + \Pr(D=1|X=x)\left(\frac{f_X(x)}{1-p} - \frac{f_X(x)}{1-p-\delta}\right) \\
&= \Pr(D=0|X=x) f_X(x)\frac{\delta}{(1-p-\delta)(1-p)} \\
&\quad - \frac{f_X(x)}{(1-p-\delta)}\left[\Pr(D_\delta=1|X=x) - \Pr(D=1|X=x)\right].
\end{aligned}
$$

Plugging this back into (A.3), we get

$$
\begin{aligned}
&(1-p)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) d\left(\frac{F_{X|D_\delta=0}(x) - F_{X|D=0}(x)}{\delta}\right) \\
&= \frac{1}{1-p-\delta}\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)\Pr(D=0|X=x) dF_X(x) \\
&- \frac{1-p}{1-p-\delta}\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)\left(\frac{\Pr(D_\delta=1|X=x) - \Pr(D=1|X=x)}{\delta}\right) dF_X(x)
\end{aligned}
$$

Under Assumption 1.3(4), we can use dominated convergence theorem to pass the limit, and conclude (pointwise in y) that

$$
\begin{aligned}
&\lim_{\delta\to 0}(1-p)\int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) d\left(\frac{F_{X|D_\delta=0}(x) - F_{X|D=0}(x)}{\delta}\right) \\
&= F_{Y|D=0}(y) - \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)\dot{P}(x) dF_X(x). \qquad \text{(A.4)}
\end{aligned}
$$

By the same arguments given for the case of (A.2), both of these results hold uniformly in $y \in \mathscr{Y}$.

Finally, we want the limit as $\delta \to 0$ of

$$\sup_{y \in \mathscr{Y}} \left| \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) dF_{X|D_\delta=1}(x) \right|$$

and

$$\sup_{y \in \mathscr{Y}} \left| \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) dF_{X|D_\delta=0}(x) \right|$$

Under Assumption 1.3(3) we have that the measures $dF_{X|D_\delta=0}$ and $dF_{X|D_\delta=1}$ converge weakly to $dF_{X|D=0}$ and $dF_{X|D=1}$ respectively. Combining this with Assumption 1.3(5), and noting that $x \mapsto F_{Y|D=d,X=x}(y)$ is bounded by definition, we have that

$$\lim_{\delta \to 0} \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) dF_{X|D_\delta=1}(x) = \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) dF_{X|D=1}(x) = F_{Y|D=1}(y).$$

and

$$\lim_{\delta \to 0} \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) dF_{X|D_\delta=0}(x) = \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) dF_{X|D=0}(x) = F_{Y|D=0}(y).$$

To make this uniform in $y \in \mathscr{Y}$, we have

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y) \left(dF_{X|D_\delta=1}(x) - dF_{X|D=1}(x) \right) \right| = 0. \tag{A.5}$$

which goes to 0 as $\delta \to 0$. Thus, we also have that

$$\sup_{y \in \mathscr{Y}} \left| \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y) \left(dF_{X|D_\delta=0}(x) - dF_{X|D=0}(x) \right) \right| \leq \int_{\mathscr{X}_0} \left| dF_{X|D_\delta=0}(x) - dF_{X|D=0}(x) \right| \tag{A.6}$$

Thus, combining (A.2), (A.4), (A.5), and (A.6), we obtain that

$$\limsup_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \frac{F_{Y_{D_\delta}}(y) - F_Y(y)}{\delta} - \int_{\mathscr{X}_1} F_{Y|D=1,X=x}(y)\dot{P}(x)dF_X(x) + \int_{\mathscr{X}_0} F_{Y|D=0,X=x}(y)\dot{P}(x)dF_X(x) \right|$$
$$= 0$$

We can write the uniform derivative as $\mathbb{E}\left[\left(F_{Y|D=1,X}(y) - F_{Y|D=0,X}(y)\right)\dot{P}(X)\right]$. Now, using the notation of Theorem 1.1, we have that

$$\dot{F}_{Y,\mathscr{D}}(y) = \mathbb{E}\left[\left(F_{Y|D=1,X}(y) - F_{Y|D=0,X}(y)\right)\dot{P}(X)\right].$$

and which is a continuous map, since $y \mapsto F_{Y|D=d,X=x}(y)$ is continuous for every $x \in \mathscr{X}_d$, and $d = 0,1$ by Assumption 1.3(5). Now, by Theorem 1.1, we have

$$\begin{aligned} M_{\tau,\mathscr{D}} &= -\frac{\dot{F}_{Y,\mathscr{D}}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))} \\ &= -\frac{\mathbb{E}\left[\left(F_{Y|D=1,X}(F_Y^{-1}(\tau)) - F_{Y|D=0,X}(F_Y^{-1}(\tau))\right)\dot{P}(X)\right]}{f_Y(F_Y^{-1}(\tau))} \end{aligned}$$

where by Assumption 1.1 $f_Y(F_Y^{-1}(\tau)) > 0$. $\square$

Proof of Lemma 1.2. Using the selection equation $D_\delta = \mathbb{1}\{U_D \le P_\delta(X)\}$, we have

$$\begin{aligned}
F_{Y(1)|D_\delta}(y|1) &= \Pr(Y(1) \le y | D_\delta = 1) \\
&= \Pr(Y(1) \le y | U_D \le P_\delta(X)) = \frac{\Pr(Y(1) \le y, U_D \le P_\delta(X))}{\Pr(U_D \le P_\delta(X))} \\
&= \frac{\int_{\mathscr{X}} \Pr(Y(1) \le y, U_D \le P_\delta(x) | X = x) f_X(x) dx}{p+\delta} \\
&= \frac{\int_{\mathscr{X}} F_{Y(1),U_D|X}(y, P_\delta(x)|x) f_X(x) dx}{p+\delta} \\
&= \frac{1}{p+\delta} \int_{\mathscr{X}} \int_{-\infty}^{y} \int_{-\infty}^{P_\delta(x)} f_{Y(1),U_D|X}(\tilde{y}, \tilde{u}|x) f_X(x) d\tilde{u} d\tilde{y} dx \\
&= \frac{1}{p+\delta} \int_{-\infty}^{y} \int_{\mathscr{X}} \left[\int_{-\infty}^{P_\delta(x)} f_{Y(1),U_D|X}(\tilde{y}, \tilde{u}|x) d\tilde{u} \right] f_X(x) dx d\tilde{y},
\end{aligned}$$

where the order of integration can be switched because the integrands are non-negative. It then follows that

$$f_{Y(1)|D_\delta}(y|1) = \frac{1}{p+\delta} \int_{\mathscr{X}} \left[\int_{-\infty}^{P_\delta(x)} f_{Y(1),U_D|X}(y, \tilde{u}|x) d\tilde{u} \right] f_X(x) dx. \tag{A.7}$$

Under Assumptions 1.5(b) and 1.5(c), we can differentiate both sides of (A.7) with respect to δ under the integral sign to get

$$\begin{aligned}
\frac{\partial f_{Y(1)|D_\delta}(y|1)}{\partial \delta} &= \frac{1}{p+\delta} \int_{\mathscr{X}} f_{Y(1),U_D|X}(y, P_\delta(x)|x) \frac{\partial P_\delta(x)}{\partial \delta} f_X(x) dx \\
&- \frac{1}{(p+\delta)^2} \int_{\mathscr{X}} \left[\int_{-\infty}^{P_\delta(x)} f_{Y(1),U_D|X}(y, \tilde{u}|x) d\tilde{u} \right] f_X(x) dx \\
&= \frac{1}{p+\delta} \int_{\mathscr{X}} f_{Y(1)|U_D,X}(y|P_\delta(x), x) \frac{\partial P_\delta(x)}{\partial \delta} f_X(x) dx - \frac{f_{Y(1)|D_\delta}(y|1)}{p+\delta} \qquad \text{(A.8)}
\end{aligned}$$

where the last line follows from (A.7).

Under Assumptions 1.5(b.i) and 1.5(c.ii), $f_{Y(1)|U_D,X}(y|P_\delta(x),x)\partial P_\delta(x)/\partial\delta$ is continuous in δ for each $y \in \mathscr{Y}(d)$ and $x \in \mathscr{X}$. In view of Assumptions 1.5(b.ii) and 1.5(c.iii), we can invoke the dominated convergence theorem to show that the map $\delta \mapsto \frac{\partial f_{Y(1)|D_\delta}(y|1)}{\partial \delta}$ is continuous for each $y \in \mathscr{Y}(d)$.

For the case of $f_{Y(0)|D_\delta}(y|0)$, we have

$$\frac{\partial f_{Y(0)|D_\delta}(y|0)}{\partial \delta} = \frac{\partial}{\partial \delta}\frac{\partial F_{Y(0)|D_\delta}(y|0)}{\partial y}.$$

Using the selection equation, we can xrite $F_{Y(0)|D_\delta}(y|0)$ as

$$\begin{aligned}
F_{Y(0)|D_\delta}(y|0) &= \Pr(Y(0) \leq y|D_\delta = 0) \\
&= \Pr(Y(0) \leq y|U_D > P_\delta(X))) \\
&= \frac{\Pr(Y(0) \leq y, U_D > P_\delta(X))}{1-p-\delta} \\
&= \frac{\Pr(Y(0) \leq y) - \Pr(Y(0) \leq y, U_D \leq P_\delta(X))}{1-p-\delta} \\
&= \frac{F_{Y(0)}(y) - \int_{\mathscr{X}} F_{Y(0),U_D|X}(y, P_\delta(x)|x) f_X(x)dx}{1-p-\delta} \\
&= \frac{1}{1-p-\delta}\left[F_{Y(0)}(y) - \int_{\mathscr{X}}\int_{-\infty}^{y}\int_{-\infty}^{P_\delta(x)} f_{Y(0),U_D|X}(\tilde{y},\tilde{u}|x) f_X(x) d\tilde{u}d\tilde{y}dx\right] \\
&= \frac{1}{1-p-\delta}\left[F_{Y(0)}(y) - \int_{-\infty}^{y}\int_{\mathscr{X}}\int_{-\infty}^{P_\delta(x)} f_{Y(0),U_D|X}(\tilde{y},\tilde{u}|x) f_X(x) d\tilde{u}dxd\tilde{y}\right],
\end{aligned}$$

where the orders of integrations can be switched because the integrands are non-negative. Therefore,

$$f_{Y(0)|D_\delta}(y|0) = \frac{1}{1-p-\delta}\left[f_{Y(0)}(y) - \int_{\mathscr{X}}\int_{-\infty}^{P_\delta(x)} f_{Y(0),U_D|X}(y,\tilde{u}|x) f_X(x) d\tilde{u}dx\right]. \quad \text{(A.9)}$$

Using Assumptions 1.5(b) and 1.5(c), we have

$$\frac{\partial f_{Y(0)|D_\delta}(y|0)}{\partial \delta} = \frac{f_{Y(0)|D_\delta}(y|0)}{1-p-\delta} - \frac{1}{1-p-\delta}\int_{\mathscr{X}} f_{Y(0),U_D|X}(y, P_\delta(x))|x)\frac{\partial P_\delta(x)}{\partial \delta} f_X(x)dx. \quad \text{(A.10)}$$

The continuity of $\delta \mapsto \frac{\partial f_{Y(0)|D_\delta}(y|0)}{\partial \delta}$ follows from the same arguments for the continuity of $\delta \mapsto \frac{\partial f_{Y(1)|D_\delta}(y|1)}{\partial \delta}$. Therefore, we have established that $\delta \mapsto f_{Y(0)|D_\delta}(y|0)$ is continuously differentiable.

□

Proof of Lemma 1.3. For any δ in N_ε, we have

$$
\begin{aligned}
F_{Y_\delta}(y) &= \Pr(Y_\delta \leq y) \\
&= \Pr((1-D_\delta)Y(0)+D_\delta Y(1) \leq y) \\
&= (p+\delta)\Pr(Y(1) \leq y|D_\delta = 1) + (1-p-\delta)\Pr(Y(0) \leq y|D_\delta = 0) \\
&= \int_{\mathscr{Y}(1)} \mathbb{1}\{\tilde{y} \leq y\}(p+\delta) f_{Y(1)|D_\delta}(\tilde{y}|1) d\tilde{y} \\
&+ \int_{\mathscr{Y}(0)} \mathbb{1}\{\tilde{y} \leq y\}(1-p-\delta) f_{Y(0)|D_\delta}(\tilde{y}|0) d\tilde{y}. \qquad \text{(A.11)}
\end{aligned}
$$

We proceed to take the first order Taylor expansion of $\delta \mapsto (p+\delta) f_{Y(1)|D_\delta}$ and $\delta \mapsto (1-p-\delta) f_{Y(0)|D_\delta}$ around $\delta = 0$, which is possible by Lemma 1.2: $f_{Y(d)|D_\delta}$ are continuously differentiable with respect to δ. We have

$$
\begin{aligned}
& (p+\delta) f_{Y(1)|D_\delta}(\tilde{y}|1) \\
= \; & p f_{Y(1)|D}(\tilde{y}|1) + \delta \cdot \left[p \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta} \bigg|_{\delta=0} + f_{Y(1)|D}(\tilde{y}|1) \right] + R(\delta; \tilde{y}, 1), \qquad \text{(A.12)}
\end{aligned}
$$

where

$$
\begin{aligned}
R(\delta;\tilde{y},1) \; &:= \; \delta \cdot \left[p \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta} \bigg|_{\delta=\tilde{\delta}_1} - p \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta} \bigg|_{\delta=0} \right] \\
&+ \; \delta \cdot \left[f_{Y(1)|D_\delta}(\tilde{y}|1) - f_{Y(1)|D_0}(\tilde{y}|1) \right] \qquad \text{(A.13)}
\end{aligned}
$$

and $0 \leq \tilde{\delta}_1 \leq \delta$. The middle point $\tilde{\delta}_1$ depends on δ. For the case of $d = 0$, we have a similar expansion.

$$
\begin{aligned}
& (1-p-\delta) f_{Y(0)|D_\delta}(\tilde{y}|0) \\
= \;& (1-p) f_{Y(0)|D}(\tilde{y}|0) + \delta \cdot \left[(1-p) \frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial \delta}\bigg|_{\delta=0} - f_{Y(0)|D}(\tilde{y}|0) \right] \\
+ \;& R(\delta; \tilde{y}, 0), \qquad \text{(A.14)}
\end{aligned}
$$

where

$$
\begin{aligned}
R(\delta; \tilde{y}, 0) \;:=\; & \delta \cdot \left[(1-p) \frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial \delta}\bigg|_{\delta=\tilde{\delta}_0} - (1-p) \frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial \delta}\bigg|_{\delta=0} \right] \\
+ \;& \delta \cdot \left[f_{Y(0)|D}(\tilde{y}|0) - f_{Y(0)|D_\delta}(\tilde{y}|0) \right] \qquad \text{(A.15)}
\end{aligned}
$$

and $0 \leq \tilde{\delta}_0 \leq \delta$. The middle point $\tilde{\delta}_0$ depends on δ.

Consider the first order derivative that appears in (A.12), when $\delta = 0$, using (A.8) we have

$$
\begin{aligned}
& \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta}\bigg|_{\delta=0} \\
= \;& \frac{1}{p} \int_{\mathscr{X}} f_{Y(1),U_D|X}(\tilde{y}, P(x)|x) \dot{P}(x) f_X(x) dx - \frac{f_{Y(1)|D}(\tilde{y}|1)}{p} \\
= \;& \frac{1}{p} \left[\int_{\mathscr{X}} f_{Y(1),U_D|X}(\tilde{y}, P(x)|x) \dot{P}(x) f_X(x) dx - f_{Y(1)|D}(\tilde{y}|1) \right] \\
= \;& \frac{1}{p} \left[\int_{\mathscr{X}} f_{Y(1)|U_D,X}(\tilde{y}|P(x), x) f_{U_D|X}\left(P(x)|x\right) \dot{P}(x) f_X(x) dx - f_{Y(1)|D}(\tilde{y}|1) \right] \\
= \;& \frac{1}{p} \left[\int_{\mathscr{X}} f_{Y(1)|U_D,X}(\tilde{y}|P(x), x) \dot{P}(x) f_X(x) dx - f_{Y(1)|D}(\tilde{y}|1) \right] \qquad \text{(A.16)}
\end{aligned}
$$

where we define

$$
\dot{P}(x) = \frac{\partial P_\delta(x)}{\partial \delta}|_{\delta=0}.
$$

Note that we have used that $U_D|X$ is uniform on $[0,1]$.

Now we substitute (A.16) in (A.12) to get

$$(p+\delta)f_{Y(1)|D_\delta}(\tilde{y}|1) = pf_{Y(1)|D}(\tilde{y}|1)+\delta\int_{\mathscr{X}} f_{Y(1)|U_D,X}(\tilde{y},P(x)|x)\dot{P}(x)f_X(x)dx + R(\delta;\tilde{y},1). \tag{A.17}$$

The first derivative in (A.14) can be handled similarly for $\delta=0$ using (A.10):

$$\left.\frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial\delta}\right|_{\delta=0} = -\frac{\int_{\mathscr{X}} f_{Y(0)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dx}{1-p}+\frac{f_{Y(0)|D}(\tilde{y}|0))}{1-p}. \tag{A.18}$$

Plugging (A.18) into (A.14), we get

$$(1-p-\delta)f_{Y(0)|D_\delta}(\tilde{y}|0) = (1-p)f_{Y(0)|D}(\tilde{y}|0)-\delta\int_{\mathscr{X}} f_{Y(0)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dx + R(\delta;\tilde{y},0). \tag{A.19}$$

Now we substitute (A.17) and (A.19) in (A.11), leading to

$$\begin{aligned} F_{Y_\delta}(y) &= \int_{\mathscr{Y}(1)} \mathbb{1}\{\tilde{y}\le y\}\left[pf_{Y(1)|D}(\tilde{y}|1)+\delta\int_{\mathscr{X}} f_{Y(1)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dx\right]d\tilde{y} \\ &+ \int_{\mathscr{Y}(0)} \mathbb{1}\{\tilde{y}\le y\}\left[(1-p)f_{Y(0)|D}(\tilde{y}|0)\right. \\ &- \left.\delta\int_{\mathscr{X}} f_{Y(0)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dx\right]d\tilde{y} \\ &+ \tilde{R}(\delta;y) \\ &= F_Y(y)+\delta\int_{\mathscr{Y}(1)}\int_{\mathscr{X}} \mathbb{1}\{\tilde{y}\le y\}f_{Y(1)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dxd\tilde{y} \\ &- \delta\int_{\mathscr{Y}(0)}\int_{\mathscr{X}} \mathbb{1}\{\tilde{y}\le y\}f_{Y(0)|U_D,X}(\tilde{y}|P(x),x)\dot{P}(x)f_X(x)dxd\tilde{y}+R_F(\delta;y) \end{aligned} \tag{A.20}$$

where the remainder $R_F(\delta;y)$ is

$$R_F(\delta;y) := \int_{\mathscr{Y}(1)} \mathbb{1}\{\tilde{y}\le y\}R(\delta;\tilde{y},1)d\tilde{y}+\int_{\mathscr{Y}(0)} \mathbb{1}\{\tilde{y}\le y\}R(\delta;\tilde{y},0)d\tilde{y}. \tag{A.21}$$

The next step is to show that the remainder in (A.21) is $o(|\delta|)$ uniformly over $y \in \mathscr{Y} = \mathscr{Y}(0) \cup \mathscr{Y}(1)$ as $\delta \to 0$, that is,

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \frac{R_F(\delta; y)}{\delta} \right| = 0.$$

Using (A.13) and (A.15), we get

$$\begin{aligned}
\sup_{y \in \mathscr{Y}} \left| \frac{R_F(\delta; y)}{\delta} \right| \leq\ & p \int_{\mathscr{Y}(1)} \left| \left. \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta} \right|_{\delta = \tilde{\delta}_1} - \left. \frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial \delta} \right|_{\delta = 0} \right| d\tilde{y} \\
& + (1-p) \int_{\mathscr{Y}(0)} \left| \left. \frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial \delta} \right|_{\delta = \tilde{\delta}_0} - \left. \frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial \delta} \right|_{\delta = 0} \right| d\tilde{y} \\
& + \int_{\mathscr{Y}(1)} \left| f_{Y(1)|D}(\tilde{y}|1) - f_{Y(1)|D_\delta}(\tilde{y}|1) \right| d\tilde{y} \\
& + \int_{\mathscr{Y}(0)} \left| f_{Y(0)|D}(\tilde{y}|0) - f_{Y(0)|D_\delta}(\tilde{y}|0) \right| d\tilde{y}.
\end{aligned}$$

Assumption 1.6 allows us to take the limit $\delta \to 0$ under the integral signs. Also, by Lemma 1.2, both $f_{Y(d)|D_\delta}(\tilde{y}|d)$ and $\frac{\partial f_{Y(d)|D_\delta}(\tilde{y}|d)}{\partial \delta}$ are continuous in δ. Therefore

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \frac{R_F(\delta; y)}{\delta} \right| = 0,$$

and we get the desired result:

$$\begin{aligned}
F_{Y_\delta}(y) &= F_Y(y) + \delta \int_{\mathscr{Y}(1)} \int_{\mathscr{X}} \mathbb{1}\{\tilde{y} \leq y\} f_{Y(1)|U_D,X}(\tilde{y}|P(x),x) \dot{P}(x) f_X(x) dx d\tilde{y} \\
&- \delta \int_{\mathscr{Y}(0)} \int_{\mathscr{X}} \mathbb{1}\{\tilde{y} \leq y\} f_{Y(0)|U_D,X}(\tilde{y}|P(x),x) \dot{P}(x) f_X(x) dx d\tilde{y} + o(|\delta|) \quad \text{(A.22)}
\end{aligned}$$

uniformly over $y \in \mathscr{Y}$ as $\delta \to 0$. This can be written more compactly as

$$F_{Y_\delta}(y) = F_Y(y) + \delta E\left[F_{Y(1)|U_D,X}(y|P(X),X)\dot{P}(X)\right] - \delta E\left[F_{Y(0)|U_D,X}(y|P(X),X)\dot{P}(X)\right] + o(|\delta|)$$

uniformly over $y \in \mathscr{Y}$ as $\delta \to 0$. □

Proof of Theorem 1.3. Using Lemma 1.3, we have

$$\begin{aligned} F_{Y_\delta}(y_{\tau,\delta}) &= F_Y(y_{\tau,\delta}) + \delta E\left[F_{Y(1)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] \\ &- \delta E\left[F_{Y(0)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] + o(|\delta|). \end{aligned}$$

Noting that $F_{Y_\delta}(y_{\tau,\delta}) = F_Y(y_\tau) = \tau$, we have

$$\begin{aligned} F_Y(y_\tau) &= F_Y(y_{\tau,\delta}) + \delta E\left[F_{Y(1)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] \\ &- \delta E\left[F_{Y(0)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] + o(|\delta|). \end{aligned} \tag{A.23}$$

Note that

$$\left|E\left[F_{Y(d)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right]\right| \le E\left|\dot{P}(X)\right| < \infty.$$

Letting $\delta \to 0$ on both sides of (A.23) yields

$$\lim_{\delta \to 0} F_Y(y_{\tau,\delta}) = F_Y(y_\tau).$$

Under the assumption that $f_Y(y_\tau) > 0$, $F_Y(\cdot)$ is continuous and strictly increasing at y_τ. Combining this with the above limit result, we conclude that $\lim_{\delta \to 0} y_{\tau,\delta} = y_\tau$.

Going back to (A.23), we have

$$\begin{aligned} &\lim_{\delta \to 0} \frac{F_Y(y_{\tau,\delta}) - F_Y(y_\tau)}{\delta} \\ = &\lim_{\delta \to 0} E\left[F_{Y(0)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] - \lim_{\delta \to 0} E\left[F_{Y(1)|U_D,X}(y_{\tau,\delta}|P(X),X)\dot{P}(X)\right] \\ = &E\left[F_{Y(0)|U_D,X}(y_\tau|P(X),X)\dot{P}(X)\right] - E\left[F_{Y(1)|U_D,X}(y_\tau|P(X),X)\dot{P}(X)\right]. \end{aligned}$$

Therefore, we have that the unconditional quantile effect is

$$\begin{aligned}
\lim_{\delta\to 0}\frac{y_{\tau,\delta}-y_\tau}{\delta} &= \frac{1}{f_Y(y_\tau)}\Big\{E\left[F_{Y(0)|U_D,X}(y_\tau|P(X),X)\dot{P}(X)\right] \\
&- E\left[F_{Y(1)|U_D,X}(y_\tau|P(X),X)\dot{P}(X)\right]\Big\}. \qquad \text{(A.24)}
\end{aligned}$$

□

Proof of Corollary 1.4. For each $d=0$ and 1, we have

$$\begin{aligned}
&\frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|U_D=P(x),X=x\right]\dot{P}(x)f_X(x)\,dx \\
=&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]\dot{P}(x)f_X(x)\,dx \\
+&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|U_D=P(x),X=x\right]\dot{P}(x)f_X(x)\,dx \\
-&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]\dot{P}(x)f_X(x)\,dx \\
=&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]f_X(x)\,dx \\
-&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]f_X(x)\left[1-\dot{P}(x)\right]dx \\
-&\ \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} \left[F_{Y(d)|D,X}(y_\tau|d,x)-F_{Y(d)|U_D,X}(y_\tau|P(x),x)\right]\dot{P}(x)f_X(x)\,dx \\
:=&\ A_\tau(d)-B_{1\tau}(d)-B_{2\tau}(d)
\end{aligned}$$

where

$$\begin{aligned}
A_\tau(d) &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]f_X(x)\,dx, \\
B_{1\tau}(d) &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\left\{Y(d)\le y_\tau\right\}|D=d,X=x\right]f_X(x)\left[1-\dot{P}(x)\right]dx, \\
B_{1\tau}(d) &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} \left[F_{Y(d)|D,X}(y_\tau|d,x)-F_{Y(d)|U_D,X}(y_\tau|P(x),x)\right]\dot{P}(x)f_X(x)\,dx.
\end{aligned}$$

So

$$\begin{aligned} M_\tau &= A_\tau(0) - B_{1\tau}(0) - B_{2\tau}(0) - [A_\tau(1) - B_{1\tau}(1) - B_{2\tau}(1)] \\ &= A_\tau - B_{1\tau} - B_{2\tau}, \end{aligned}$$

where

$$\begin{aligned} A_\tau &= A_\tau(0) - A_\tau(1) \\ &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\{Y \le y_\tau\} | D = 0, X = x\right] f_X(x)\, dx \\ &- \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} E\left[\mathbb{1}\{Y \le y_\tau\} | D = 1, X = x\right] f_X(x)\, dx, \end{aligned}$$

$$\begin{aligned} B_{1\tau} &= B_{1\tau}(0) - B_{1\tau}(1) \\ &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[F_{Y(0)|D,X}(y_\tau|0,x) - F_{Y(1)|D,X}(y_\tau|1,x)\right] \left[1 - \dot{P}(x)\right] f_X(x)\, dx, \end{aligned}$$

and

$$\begin{aligned} B_{2\tau} &= B_{2\tau}(0) - B_{2\tau}(1) \\ &= \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[F_{Y(0)|D,X}(y_\tau|0,x) - F_{Y(0)|U_D,X}(y_\tau|P(x),x)\right] \dot{P}(x) f_X(x)\, dx \\ &+ \frac{1}{f_Y(y_\tau)} \int_{\mathscr{X}} \left[F_{Y(1)|U_D,X}(y_\tau|P(x),x) - F_{Y(1)|D,X}(y_\tau|1,x)\right] \dot{P}(x) f_X(x)\, dx. \end{aligned}$$

□

Proof of Lemma 1.4. For a given δ, $s_\mu(\delta)$ satisfies $\Pr(D_\delta = 1) = p + \delta$. But

$$\Pr(D_\delta = 1) = E\left[P_\delta(X)\right] = \int_{\mathscr{X}} F_{V|X}(\mu(x) + s_\mu(\delta)|x) f_X(x) dx,$$

and so

$$p+\delta=\int_{\mathscr{X}}F_{V|X}(\mu(x)+s_{\mu}(\delta)|x)f_X(x)dx. \tag{A.25}$$

Note that $s_\mu(0)=0$. We need to find the derivative of the implicit function $s_\mu(\delta)$ with respect to δ. Define

$$t(\delta,s)=p+\delta-\int_{\mathscr{X}}F_{V|X}(\mu(x)+s|x)f_X(x)dx. \tag{A.26}$$

By Theorem 9.28 in Rudin (1976), we need to show that t is continuously differentiable in a neighborhood around $(0,0)$ of (δ,s). We do this, by showing that the partial derivatives of (A.26) with respect to δ and s exist and are continuous (See Theorem 9.21 in Rudin (1976)).

For the partial derivative with respect to δ, we have $\partial t(\delta,s)/\partial\delta=1$, which is obviously continuous in (δ,s). For the partial derivative with respect to s, we use Assumption (iii) in the lemma to obtain

$$\frac{\partial t(\delta,s)}{\partial s}=-\int_{\mathscr{X}}f_{V|X}(\mu(x)+s|x)f_X(x)dx.$$

The function is trivially continuous in δ. In view of the continuity of $f_{V|X}(v|x)$ in v for almost all x, the dominated convergence theorem implies that $\partial t(\delta,s)/\partial s$ is also continuous in s. Therefore, we can apply the implicit function theorem to obtain $s'_\mu(\delta)$ in a neighborhood of $\delta=0$. Taking the derivative of (A.25) with respect to δ, we get

$$\frac{\partial s_\mu(\delta)}{\partial\delta}=s'_\mu(\delta)=\frac{1}{\int_{\mathscr{X}}f_{V|X}(\mu(x)+s_\mu(\delta)|x)f_X(x)dx}.$$

In particular, for $\delta=0$, we have

$$s'_\mu(0)=\frac{1}{\int_{\mathscr{X}}f_{V|X}(\mu(x)|x)f_X(x)dx}.$$

For the propensity score, we have

$$P^{\mu}_{\delta}(x)=F_{V|X}(\mu(x)+s_\mu(\delta)|x)$$

and so

$$\begin{aligned}\frac{\partial P^{\mu}_{\delta}(x)}{\partial\delta} &= f_{V|X}(\mu(x)+s_{\mu}(\delta)|x)\frac{\partial s_{\mu}(\delta)}{\partial\delta}\\ &= \frac{f_{V|X}(\mu(x)+s_{\mu}(\delta)|x)}{\int_{\mathscr{X}} f_{V|X}(\mu(x)+s_{\mu}(\delta)|x)f_X(x)dx}.\end{aligned}$$

Evaluating the above at $\delta=0$ gives us

$$\dot{P}^{\mu}(x)=\frac{f_{V|X}(\mu(x)|x)}{\int_{\mathscr{X}} f_{V|X}(\mu(x)|x)f_X(x)dx}.$$

□

Proof of Corollary 1.6. It follows from Theorem 1.3 that

$$\begin{aligned}M_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(0)\le y_\tau\}|U_D=P(x),X=x\right]f_X(x)\,dx,\\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(1)\le y_\tau\}|U_D=P(x),X=x\right]f_X(x)\,dx. \qquad \text{(A.27)}\end{aligned}$$

Since the propensity score does not depend on x because $\mu(\cdot)=\mu_0$, a constant, we have $U_D=P(x)=p$, where $p:=F_V(\mu_0)=\Pr(D=1)$. But U_D is independent of X, so

$$\begin{aligned}M_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(0)\le y_\tau\}|U_D=p,X=x\right]f_{X|U_D}(x|p)\,dx\\ &- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{X}} E\left[\mathbb{1}\{Y(1)\le y_\tau\}|U_D=p,X=x\right]f_{X|U_D}(x|p)\,dx\\ &= \frac{1}{f_Y(y_\tau)}E\left[\mathbb{1}\{Y(0)\le y_\tau\}-\mathbb{1}\{Y(1)\le y_\tau\}|U_D=p\right]. \qquad \text{(A.28)}\end{aligned}$$

In the decomposition $M_{\tau,\mu}=A_{\tau,\mu}-B_{\tau,\mu}$, the formula for $A_{\tau,\mu}$ follows from the same argument as above, and the formula for $B_{\tau,\mu}$ follows from Corollary 1.4 because $B_{1\tau}=0$ and $B_{2\tau}$ simplifies to the given expression. □

Proof of Lemma 1.5. We deal with each term at a time of (1.16). We have

$$\begin{aligned}\mathscr{I}_1 &= \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\}\, f_{\tilde{U}|X=x-\delta}(\tilde{u})\left[f_X(x-\delta) - f_X(x)\right] d\tilde{u}dx \\ &= -\delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\}\, f_{\tilde{U}|X=x-\delta}(\tilde{u}) f'_X(x) d\tilde{u}dx + R_1(\delta,y)\end{aligned}$$

where $f'_X(x)$ is the derivative of the density and the remainder $R(\delta,y)$ is

$$\begin{aligned}R_1(\delta,y) := -\delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} &\mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\}\, f_{\tilde{U}|X=x-\delta}(\tilde{u}) \\ &\times \left[f'_X(x-\tilde{\delta}) - f'_X(x)\right] d\tilde{u}dx\end{aligned}$$

where $0 \leq \tilde{\delta} \leq \delta$, and $\tilde{\delta}$ depends on x.

Now we show that $\sup_{y\in\mathscr{Y}} |R(\delta,y)| = o(|\delta|)$ as $\delta \to 0$.

$$\begin{aligned}R_1(\delta,y) = &-\delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\}\, f_{\tilde{U}|X=x-\delta}(\tilde{u}) f'_X(x-\tilde{\delta}) d\tilde{u}dx \\ &+ \delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\}\, f_{\tilde{U}|X=x}(\tilde{u}) f'_X(x) d\tilde{u}dx \\ &+ \delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\}\,\mathbb{1}\{x_l \leq x-\delta \leq x_u\})\left[f_{\tilde{U}|X=x-\delta}(\tilde{u}) - f_{\tilde{U}|X=x}(\tilde{u})\right] f'_X(x) d\tilde{u}dx.\end{aligned}$$

The first term and the second term, leaving aside the factor δ have the same limit uniformly in $y \in \mathscr{Y}$ because both indicator functions are bounded, and because of the domination assumption (1.17). For the third term, we use (1.17) to pass the limit, so that it is also $o(|\delta|)$ uniformly in $y \in \mathscr{Y}$. Hence,

$$\sup_{y\in\mathscr{Y}} R_1(\delta,y) = o(|\delta|).$$

Therefore, we can write

$$\frac{\mathscr{I}_1(y)}{\delta} = -\int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f'_X(x) d\tilde{u}dx + o(1)$$

where the $o(1)$ term does not depend on y. Now we need to pass the limit as $\delta \to 0$ inside the integral. Again, we can use (1.17). We have that

$$\begin{aligned}
&\lim_{\delta\to 0} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f'_X(x) \\
&= \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x \leq x_u\} f_{\tilde{U}|X=x}(\tilde{u}) f'_X(x)
\end{aligned}$$

and the sequence is dominated by

$$\begin{aligned}
\left| \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_{\tilde{U}|X=x-\delta}(\tilde{u}) f'_X(x) \right| &\leq f_{\tilde{U}|X=x-\delta}(\tilde{u}) \left| f'_X(x) \right| \\
&= \sup_{\delta\in N_\varepsilon} f_{\tilde{U}|X=x-\delta}(\tilde{u}) \sup_{\delta'\in N_\varepsilon} |f'_X(x-\delta')|
\end{aligned}$$

which is integrable by Assumption 1.11. Therefore, we have

$$\lim_{\delta\to 0} \sup_{y\in\mathscr{Y}} \left| \frac{\mathscr{I}_1(y)}{\delta} + \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x \leq x_u\} f_{\tilde{U}|X=x}(\tilde{u}) f'_X(x) d\tilde{u}dx \right| = 0.$$

which can be written as

$$\lim_{\delta\to 0} \sup_{y\in\mathscr{Y}} \left| \frac{\mathscr{I}_1(y)}{\delta} + \int_{x_l}^{x_u}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} f_{\tilde{U}|X=x}(\tilde{u}) f'_X(x) d\tilde{u}dx \right| = 0. \tag{A.29}$$

Now, for the next term, $\mathscr{I}_2(y)$, using Assumption 1.10, we do a Taylor expansion to obtain

$$\mathscr{I}_2(y) = -\delta \int_{\mathbb{R}}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \leq y\} \mathbb{1}\{x_l \leq x-\delta \leq x_u\} f_X(x) f'_{\tilde{U}|X=x}(\tilde{u}) d\tilde{u}dx + R_2(y,\delta)$$

where the remainder is

$$R_2(y,\delta) := -\delta \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} \mathbb{1}\{x_l \le x-\delta \le x_u\} f_X(x) \\ \times \left[f'_{\tilde{U}|X=x-\tilde{\delta}}(\tilde{u}) - f'_{\tilde{U}|X=x}(\tilde{u}) \right] d\tilde{u}dx$$

for $\tilde{\delta}$ such that $0 \le \tilde{\delta} \le \delta$, and it depends on both x and $\tilde{u}$. Under equation (1.18) in Assumption 1.10, $R_2(y,\delta) = o(|\delta|)$, uniformly in $y \in \mathscr{Y}$. Thus we have that

$$\lim_{\delta \to 0} \sup_{y \in \mathscr{Y}} \left| \frac{\mathscr{I}_2(y)}{\delta} + \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} \mathbb{1}\{x_l \le x \le x_u\} f_X(x) f'_{\tilde{U}|X=x}(\tilde{u}) d\tilde{u}dx \right| = 0 \quad \text{(A.30)}$$

with no domination required since the only function of δ that remains is $\mathbb{1}\{x_l \le x-\delta \le x_u\}$.

The last term is

$$\mathscr{I}_3(y) := \int_{\mathbb{R}} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} \left[\mathbb{1}\{x_l \le x-\delta \le x_u\} - \mathbb{1}\{x_l \le x \le x_u\} \right] f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx,$$

which captures the discrepancy in the supports of the target variable. While it converges to 0 as $\delta \to 0$, we are actually interested in $\mathscr{I}_3(y)/\delta$ as $\delta \to 0$. For a positive δ we have

$$\mathbb{1}\{x_l \le x-\delta \le x_u\} - \mathbb{1}\{x_l \le x \le x_u\} = \begin{cases} -1 & \text{if } x \in [x_l, x_l+\delta) \\ 0 & \text{if } x \in [x_l+\delta, x_u] \\ 1 & \text{if } x \in (x_u, x_u+\delta] \end{cases}$$

Thus, we can write $\mathscr{I}_3(y)$ as

$$\mathscr{I}_3(y) := \int_{x_u}^{x_u+\delta} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx \\ - \int_{x_l}^{x_l+\delta} \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u}) \le y\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x) d\tilde{u}dx$$

By the Fundamental Theorem of Calculus, we have that

$$\lim_{\delta\to 0}\frac{\mathscr{I}_3(y)}{\delta} := \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_u,\tilde{u})\le y\} f_{\tilde{U}|X=x_u}(\tilde{u}) f_X(x_u)d\tilde{u} \tag{A.31}$$

$$- \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_l,\tilde{u})\le y\} f_{\tilde{U}|X=x_l}(\tilde{u}) f_X(x_l)d\tilde{u}$$

$$= f_X(x_u)F_{Y|X=x_u}(y) - f_X(x_l)F_{Y|X=x_l}(y)$$

$$= 0. \tag{A.32}$$

by Assumption 1.7 because $f_X(x_l) = f_X(x_u) = 0$.[3] Thus, using equations (A.29), (A.30), and (A.31) we have

$$\lim_{\delta\to 0}\sup_{y\in\mathscr{Y}}\left|\frac{F_{Y^*}(y)-F_Y(y)}{\delta} - \dot{F}(y)\right| = 0,$$

where

$$\dot{F}_Y(y) = -\int_{x_l}^{x_u}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u})\le y\} f_{\tilde{U}|X=x}(\tilde{u}) f'_X(x)d\tilde{u}dx$$
$$-\int_{x_l}^{x_u}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u})\le y\} f_X(x) f'_{\tilde{U}|X=x}(\tilde{u})d\tilde{u}dx$$

Since $Y = r(X,\tilde{U})$, we can write $\dot{F}_Y(y)$ as

$$\dot{F}_Y(y) = -\int_{x_l}^{x_u} F_{Y|X=x}(y)f'_X(x)dx - \int_{x_l}^{x_u}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u})\le y\}\frac{f'_{\tilde{U}|X=x}(\tilde{u})}{f_{\tilde{U}|X=x}(\tilde{u})} f_{X,\tilde{U}}(x,\tilde{u})d\tilde{u}dx$$

[3]Note that if we do not assume that $f_X(x) = 0$ on the boundary of its support, *i.e.*, at x_l and x_u, convergence may not hold uniformly since:

$$\left|\mathscr{I}_3(y) - \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_u,\tilde{u})\le y\} f_{\tilde{U}|X=x_u}(\tilde{u}) f_X(x_u)d\tilde{u} - \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_l,\tilde{u})\le y\} f_{\tilde{U}|X=x_l}(\tilde{u}) f_X(x_l)d\tilde{u}\right|$$
$$\le \left|\int_{x_u}^{x_u+\delta}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u})\le y\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x)d\tilde{u}dx - \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_u,\tilde{u})\le y\} f_{\tilde{U}|X=x_u}(\tilde{u}) f_X(x_u)d\tilde{u}\right|$$
$$+\left|\int_{x_l}^{x_l+\delta}\int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x,\tilde{u})\le y\} f_{\tilde{U}|X=x}(\tilde{u}) f_X(x)d\tilde{u}dx - \int_{\tilde{\mathscr{U}}} \mathbb{1}\{r(x_l,\tilde{u})\le y\} f_{\tilde{U}|X=x_l}(\tilde{u}) f_X(x_l)d\tilde{u}\right|$$

and that bound will likely depend on y.

Using Assumption 1.7, we can write the first term on the right hand side as an average derivative

$$-\int_{x_l}^{x_u} F_{Y|X=x}(y) f_X'(x) dx = \int_{x_l}^{x_u} \frac{\partial F_{Y|X=x}(y)}{\partial x} f_X(x) dx = \mathbb{E}\left[\left. \frac{\partial F_{Y|X=x}(y)}{\partial x} \right|_{x=X} \right],$$

which is what Firpo et al. (2009) obtain. The second term of $\dot{F}_Y(y)$, will be the bias term. We have

$$\begin{aligned} \frac{f'_{\tilde{U}|X=x}(\tilde{u})}{f_{\tilde{U}|X=x}(\tilde{u})} &= \frac{\partial f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} \frac{1}{f_{\tilde{U}|X=x}(\tilde{u})} \\ &= \frac{\partial \log f_{\tilde{U}|X=x}(\tilde{u})}{\partial x} = \frac{\partial}{\partial x} \log \frac{f_{\tilde{U}|X=x}(\tilde{u})}{f_{\tilde{U}}(\tilde{u})} = \frac{\partial}{\partial x} \log \frac{f_{X,\tilde{U}}(x,\tilde{u})}{f_X(x) f_{\tilde{U}}(\tilde{u})} \end{aligned}$$

So, we write

$$\dot{F}_Y(y) = \mathbb{E}\left[\left. \frac{\partial F_{Y|X=x}(y)}{\partial x} \right|_{x=X} \right] - \mathbb{E}\left[\mathbb{1}\left\{ r(X,\tilde{U}) \leq y \right\} \frac{\partial}{\partial x} \log \left. \frac{f_{X,\tilde{U}}(x,\tilde{u})}{f_X(x) f_{\tilde{U}}(\tilde{u})} \right|_{x=X,\tilde{u}=\tilde{U}} \right].$$

□

Proof of Theorem 1.8. Under $X, W \perp U$, we have a simpler expression for $\mathscr{I}_x(x,w,u)$. Indeed

$$\begin{aligned} \mathscr{I}_x(x,w,u) &:= \frac{\partial}{\partial x} \log \frac{f_{X,W,U}(x,w,u)}{f_X(x) f_W(w) f_U(u)} \\ &= \frac{\partial}{\partial x} \log \frac{f_{X,W}(x,w)}{f_X(x) f_W(w)} \\ &=: \mathscr{I}_x(x,w) \end{aligned}$$

thus we can ignore U. The bias is then

$$\begin{aligned} B_\tau &:= \frac{1}{f_Y(y_\tau)} \mathbb{E}\left[\mathbb{1}\left\{ r(X,W,U) \leq y_\tau \right\} \mathscr{I}_x(X,W) \right] \\ &= \frac{1}{f_Y(y_\tau)} \mathbb{E}\left[F_{Y|X,W}(y_\tau) \mathscr{I}_x(X,W) \right], \end{aligned}$$

Thus the bias can be computed from the data. Alternatively, let's analyze the bias more closely. Under Assumption 1.7

$$
\begin{aligned}
B_\tau &= \frac{1}{f_Y(y_\tau)} \mathbb{E}\left[F_{Y|X,W}(y_\tau) \frac{\partial}{\partial x} \log \frac{f_{X,W}(x,w)}{f_X(x) f_W(w)} \bigg|_{x=X, w=W} \right] \\
&= \frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau) \frac{\partial f_{W|X=x}(x,w)}{\partial x} \frac{1}{f_{W|X=x}(x,w)} f_{X,W}(x,w) dx dw \\
&= \frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau) \frac{\partial f_{W|X=x}(x,w)}{\partial x} f_X(x) dx dw \\
&= -\frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} \frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} f_{W|X=x}(x,w) f_X(x) dx dw \\
&- \frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau) f_{W|X=x}(x,w) f_X'(x) dx dw \\
&= -\frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} \frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} f_{W|X=x}(x,w) f_X(x) dx dw \\
&- \frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} F_{Y|X=x}(y_\tau) f_X'(x) dx \\
&= -\frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} \frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} f_{W|X=x}(x,w) f_X(x) dx dw \\
&+ \frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \frac{\partial F_{Y|X=x}(y_\tau)}{\partial x} f_X(x) dx \\
&= -\frac{1}{f_Y(y_\tau)} \int_{x_l}^{x_u} \int_{\mathscr{W}} \frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} f_{W|X=x}(x,w) f_X(x) dx dw \\
&- A_\tau.
\end{aligned}
$$

Now, suppose that $X \perp U | W$. Then, we have $f_{X,U|W=w}(x,u) = f_{X|W=w}(x) f_{U|W=w}(u)$. So that

$$
\begin{aligned}
\mathscr{I}_x(x,w,u) &:= \frac{\partial}{\partial x}\log\frac{f_{X,W,U}(x,w,u)}{f_X(x)f_W(w)f_U(u)} \\
&= \frac{\partial}{\partial x}\log\frac{f_{X,U|W=w}(x,w,u)}{f_X(x)f_U(u)} \\
&= \frac{\partial}{\partial x}\log\frac{f_{X|W=w}(x)f_{U|W=w}(u)}{f_X(x)f_U(u)} \\
&= \frac{\partial}{\partial x}\log\frac{f_{X|W=w}(x)}{f_X(x)} \\
&= \frac{\partial \log f_{W|X=x}(w)}{\partial x}
\end{aligned}
$$

So the bias is now (using the law of iterated expectations)

$$
\begin{aligned}
B_\tau &:= \frac{1}{f_Y(y_\tau)}\mathbb{E}\left[\mathbb{1}\left\{r(X,W,U)\le y_\tau\right\}\frac{\partial \log f_{W|X=x}(w)}{\partial x}\bigg|_{x=X,w=W}\right] \\
&= \frac{1}{f_Y(y_\tau)}\mathbb{E}\left[F_{Y|X,W}(y_\tau)\frac{\partial \log f_{W|X=x}(w)}{\partial x}\bigg|_{x=X,w=W}\right] \\
&= \frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u}\int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau)\frac{\partial \log f_{W|X=x}(w)}{\partial x}f_{X,W}(x,w)dwdx \\
&= \frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u}\int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau)\frac{\partial f_{W|X=x}(w)}{\partial x}\frac{f_{X,W}(x,w)}{f_{W|X=x}(w)}dwdx \\
&= \frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u}\int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau)\frac{\partial f_{W|X=x}(w)}{\partial x}f_X(x)dwdx
\end{aligned}
$$

Under Assumption 1.7, we do integration by parts to obtain that the bias is 0.

$$
\begin{aligned}
B_\tau :=& -\frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u}\int_{\mathscr{W}} F_{Y|X=x,W=w}(y_\tau) f_X'(x) f_{W|X=x}(w)dwdx \\
& -\frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u}\int_{\mathscr{W}} \frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x} f_X(x) f_{W|X=x}(w)dwdx \\
&= -\frac{1}{f_Y(y_\tau)}\int_{x_l}^{x_u} F_{Y|X=x}(y_\tau) f_X'(x)dx \\
& -\frac{1}{f_Y(y_\tau)}\mathbb{E}\left[\left.\frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x}\right|_{x=X,w=W}\right] \\
&= -A_\tau - \frac{1}{f_Y(y_\tau)}\mathbb{E}\left[\left.\frac{\partial F_{Y|X=x,W=w}(y_\tau)}{\partial x}\right|_{x=X,w=W}\right]
\end{aligned}
$$

We get the same result as when $X,W \perp U$. □

Proof of Theorem 2.1. Follows directly from an application of Assumption 2.1(b) to Corollary 1.5. □

Proof of Lemma 2.1. For a given δ, $s_z(\delta)$ satisfies $\Pr(D_\delta = 1) = p + \delta$. But

$$\Pr(D_\delta = 1) = E\left[P_\delta(W)\right] = \int_{\mathscr{W}} F_{V|W}(\mu(z + g(w)s_z(\delta), x)|w) f_W(w)dw,$$

and so

$$p + \delta = \int_{\mathscr{W}} F_{V|W}(\mu(z + g(w)s_z(\delta), x)|w) f_W(w)dw. \tag{A.33}$$

Note that $s_z(0) = 0$. We need to find the derivative of the implicit function $s_z(\delta)$ with respect to δ. Define

$$t(\delta, s) = p + \delta - \int_{\mathscr{W}} F_{V|W}(\mu(z + g(w)s, x)|w) f_W(w)dw. \tag{A.34}$$

By Theorem 9.28 in Rudin (1976), we need to show that t is continuously differentiable in a neighborhood around $(0,0)$ of (δ,s). We do this, by showing that the partial derivatives of (A.34) with respect to δ and s exist and are continuous (See Theorem 9.21 in Rudin (1976)).

For the partial derivative with respect to δ, we have $\partial t(\delta,s)/\partial\delta = 1$, which is obviously continuous in (δ,s). For the partial derivative with respect to s, we use Assumption (iii) in the Lemma to obtain

$$\frac{\partial t(\delta,s)}{\partial s} = -\int_{\mathscr{W}} f_{V|W}(\mu(z+g(w)s,x)|w)\mu_z'(z+g(w)s,x)\,g(w)f_W(w)dw.$$

The function is trivially continuous in δ. In view of the continuity of $f_{V|W}(v|w)$ in v for almost all w, the dominated convergence theorem implies that $\partial t(\delta,s)/\partial s$ is also continuous in s. Therefore, we can apply the implicit function theorem to obtain $s_z'(\delta)$ in a neighborhood of $\delta = 0$. Taking the derivative of (A.33) with respect to δ, we get

$$\begin{aligned}\frac{\partial s_z(\delta)}{\partial\delta} &= \frac{1}{\int_{\mathscr{W}} f_{V|W}\left(\mu\left(z+g\left(w\right)s_z\left(\delta\right),x\right)|w\right)\mu_z'\left(z+g\left(w\right)s_z\left(\delta\right),x\right)g(w)f_W(w)dw} \\ &= \frac{1}{E\left[f_{V|W}\left(\mu\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)|W\right)\mu_z'\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)g(W)\right]}.\end{aligned}$$

Next, we have

$$P_\delta(z,x) = \Pr(D_\delta = 1|Z=z, X=x) = F_{V|W}\left(\mu(z+g(w)s_z(\delta),x)|w\right).$$

So

$$\begin{aligned}\frac{\partial P_\delta(z,x)}{\partial\delta} &= f_{V|W}\left(\mu(z+g(w)s_z(\delta),x)|w\right)\mu_z'(z+g(w)s_z(\delta),x)\,g(w)\frac{\partial s_z(\delta)}{\partial\delta} \\ &= \frac{f_{V|W}\left(\mu(z+g(w)s_z(\delta),x)|w\right)\mu_z'(z+g(w)s_z(\delta),x)\,g(w)}{E\left[f_{V|W}\left(\mu\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)|W\right)\mu_z'\left(Z+g\left(W\right)s_z\left(\delta\right),X\right)g(W)\right]}.\end{aligned}$$

It then follows that

$$\left.\frac{\partial s_z(\delta)}{\partial \delta}\right|_{\delta=0} = \frac{1}{E\left[f_{V|W}\left(\mu\left(W\right)|W\right)\mu_z'\left(W\right)g(W)\right]},$$

and

$$\left.\frac{\partial P_\delta(z,x)}{\partial \delta}\right|_{\delta=0} = \frac{f_{V|W}\left(\mu(w)|w\right)\mu_z'(w)\,g(w)}{E\left[f_{V|W}\left(\mu\left(W\right)|W\right)\mu_z'\left(W\right)g(W)\right]}.$$

□

Proof of Theorem 2.2. The theorem follows directly from an application of Lemma 2.1 to Theorem 1.3. □

Proof of Corollary 2.4. It is easy to see that

$$\begin{aligned}
& \int_{\mathscr{Y}} yE\left[\left\{f_{Y(1)|U_D,W}(y|P\left(W\right),W) - f_{Y(0)|U_D,W}(y|P\left(W\right),W)\right\}\dot{P}\left(W\right)\right]dy \\
= \quad & E\left\{\left[Y(1) - Y\left(0\right)|U_D = P(W),W\right]\dot{P}(W)\right\}.
\end{aligned}$$

Hence it suffices to show that $\lim_{\delta\to 0}\delta^{-1}\int_{\mathscr{Y}} ydR_F(\delta;y) = 0$. Under Assumption 2.2, we have

$$\begin{aligned}
\left|\frac{1}{\delta}\int_{\mathscr{Y}} ydR_F(\delta;y)\right| \le {} & p\int_{\mathscr{Y}(1)}|\tilde{y}|\cdot\left|\left.\frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial\delta}\right|_{\delta=\tilde{\delta}_1} - \left.\frac{\partial f_{Y(1)|D_\delta}(\tilde{y}|1)}{\partial\delta}\right|_{\delta=0}\right|d\tilde{y} \\
& + (1-p)\int_{\mathscr{Y}(0)}|\tilde{y}|\cdot\left|\left.\frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial\delta}\right|_{\delta=\tilde{\delta}_0} - \left.\frac{\partial f_{Y(0)|D_\delta}(\tilde{y}|0)}{\partial\delta}\right|_{\delta=0}\right|d\tilde{y} \\
& + \int_{\mathscr{Y}(1)}|\tilde{y}|\cdot\left|f_{Y(1)|D}(\tilde{y}|1) - f_{Y(1)|D_\delta}(\tilde{y}|1)\right|d\tilde{y} \\
& + \int_{\mathscr{Y}(0)}|\tilde{y}|\cdot\left|f_{Y(0)|D}(\tilde{y}|0) - f_{Y(0)|D_\delta}(\tilde{y}|0)\right|d\tilde{y}.
\end{aligned}$$

As in the proof of Lemma 1.3, each term in the above upper bound converges to zero as $\delta\to 0$.

Therefore,

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{\mathscr{Y}} y dR_F(\delta; y) = 0,$$

as desired. □

Proof of Lemma 2.3. We have

$$\begin{aligned}
\hat{f}_Y(y) - f_Y(y) &= \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{Y_i - y}{h}\right) - f_Y(y) \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{Y_i - y}{h}\right) - E\hat{f}_Y(y) + E\hat{f}_Y(y) - f_Y(y) \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{Y_i - y}{h}\right) - E\hat{f}_Y(y) + B_{f_Y}(y) + o_p(h^2),
\end{aligned}$$

where

$$B_{f_Y}(y) = \frac{1}{2} h^2 f_Y''^2(y) \int_{-\infty}^{\infty} u^2 K(u) du.$$

We write this concisely as

$$\hat{f}_Y(y) - f_Y(y) = \frac{1}{n} \sum_{i=1}^{n} \psi_{f_Y,i}(y,h) + B_{f_Y}(y) + o_p(h^2),$$

where

$$\psi_{f_Y,i}(y,h) := \frac{1}{h} K\left(\frac{Y_i - y}{h}\right) - E\frac{1}{h} K\left(\frac{Y_i - y}{h}\right) = O_p(n^{-1/2} h^{-1/2}).$$

Since $K(u)$ is twice continuously differentiable, we use a Taylor expansion to obtain

$$\hat{f}_Y(\hat{y}_\tau) - \hat{f}_Y(y_\tau) = \hat{f}_Y'(y_\tau)(\hat{y}_\tau - y_\tau) + \frac{1}{2} \hat{f}''(\tilde{y}_\tau)(\hat{y}_\tau - y_\tau)^2 \tag{A.35}$$

for some $\tilde{y}_\tau$ between $\hat{y}_\tau$ and y_τ. The first and second derivatives are

$$\hat{f}_Y'(y) = -\frac{1}{nh^2} \sum_{i=1}^{n} K'\left(\frac{Y_i - y}{h}\right), \quad \hat{f}_Y''(y) = \frac{1}{nh^3} \sum_{i=1}^{n} K''\left(\frac{Y_i - y}{h}\right).$$

To find the order of $\hat{f}_Y''(y)$, we calculate its mean and variance. We have

$$\begin{aligned} E\left[\hat{f}_Y''(y)\right] &= \frac{1}{nh^3}\sum_{i=1}^{n} K''\left(\frac{Y_i-y}{h}\right) = O\left(\frac{1}{h^2}\right), \\ var\left[\hat{f}_Y''(y)\right] &\leq \frac{n}{(nh^3)^2} E\left[K''\left(\frac{Y_i-y}{h}\right)\right]^2 = O\left(\frac{1}{nh^5}\right). \end{aligned}$$

Therefore, when $nh^3 \to \infty$,

$$\hat{f}_Y''(y) = O_p\left(h^{-2}\right),$$

for any y. That is, for any $\varepsilon > 0$, there exists an $M > 0$ such that

$$\Pr\left(h^2\left|\hat{f}_Y''(y_\tau)\right| > \frac{M}{2}\right) < \frac{\varepsilon}{2}$$

when n is large enough.

Suppose we choose M so large that we also have

$$\Pr\left(\sqrt{n}\left|\tilde{y}_\tau - y_\tau\right| > M\right) < \frac{\varepsilon}{2}$$

when n is large enough. Then, when n is large enough,

$$\begin{aligned} \Pr\left(h^2\hat{f}_Y''(\tilde{y}_\tau) > \frac{M}{2}\right) &\leq \Pr\left(h^2\left|\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right| > \frac{M}{2}\right) + \Pr\left(h^2\left|\hat{f}_Y''(y_\tau)\right| > \frac{M}{2}\right) \\ &\leq \Pr\left(h^2\left[\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right] > \frac{M}{2}\right) + \frac{\varepsilon}{2} \\ &\leq \Pr\left(h^2\left[\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right] > \frac{M}{2}, \sqrt{n}\left|\tilde{y}_\tau - y_\tau\right| < M\right) + \varepsilon. \quad \text{(A.36)} \end{aligned}$$

When $\sqrt{n}\left|\tilde{y}_\tau - y_\tau\right| < \sqrt{h}$, we have

$$\begin{aligned} h^2\left|\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right| &\leq \frac{1}{nh}\sum_{i=1}^{n}\left|K''\left(\frac{Y_i - y_\tau}{h}\right) - K''\left(\frac{Y_i - \tilde{y}_\tau}{h}\right)\right| \\ &\leq L_K \cdot \frac{1}{h^2}\frac{\sqrt{h}M}{\sqrt{n}} = L_K \cdot \frac{M}{\sqrt{nh^3}} \end{aligned}$$

by the Lipschitz continuity of $K''(\cdot)$ with Lipschitz constant L_K. When $\sqrt{h} \le \sqrt{n}\,|\tilde{y}_\tau - y_\tau| < M$, we have

$$\left|\frac{Y_i - y_\tau}{h} - \frac{Y_i - \tilde{y}_\tau}{h}\right| = \frac{\sqrt{n}\,|\tilde{y}_\tau - y_\tau|}{h} > \frac{1}{\sqrt{h}} \to \infty.$$

Using the second condition on $K''(\cdot)$, we have, for $\sqrt{h} \le \sqrt{n}\,|\tilde{y}_\tau - y_\tau| < M$,

$$\left|K''\left(\frac{Y_i - y_\tau}{h}\right) - K''\left(\frac{Y_i - \tilde{y}_\tau}{h}\right)\right| \le C_2 \frac{M}{nh^2},$$

and

$$\begin{aligned} h^2\left|\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right| &\le \frac{1}{nh}\sum_{i=1}^{n}\left|K''\left(\frac{Y_i - y_\tau}{h}\right) - K''\left(\frac{Y_i - \tilde{y}_\tau}{h}\right)\right| \\ &\le C_2\frac{1}{nh^3} = O\left(\frac{1}{\sqrt{nh^3}}\right). \end{aligned}$$

Hence, in both cases, $h^2\left|\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right| = O_p\left(n^{-1/2}h^{-3/2}\right)$. As a result,

$$\Pr\left(h^2\left[\hat{f}_Y''(\tilde{y}_\tau) - \hat{f}_Y''(y_\tau)\right] > \frac{M}{2}, \sqrt{n}\,|\tilde{y}_\tau - y_\tau| < M\right) \to 0.$$

Combining this with (A.36), we obtain

$$h^2\hat{f}_Y''(\tilde{y}_\tau) = O_p(1) \text{ and } \hat{f}_Y''(\tilde{y}_\tau)(\hat{y}_\tau - y_\tau)^2 = O_p\left(n^{-1}h^{-2}\right).$$

In view of (A.35), we then have

$$\hat{f}_Y(\hat{y}_\tau) - \hat{f}_Y(y_\tau) = \hat{f}_Y'(y_\tau)(\hat{y}_\tau - y_\tau) + O_p(n^{-1}h^{-2}).$$

Now, using Theorem 2.2, we can write

$$\begin{aligned}
\hat{f}_Y(\hat{y}_\tau) - \hat{f}_Y(y_\tau) &= f'_Y(y_\tau)(\hat{y}_\tau - y_\tau) + \left[\hat{f}'_Y(y_\tau) - f'_Y(y_\tau)\right](\hat{y}_\tau - y_\tau) + O_p(n^{-1}h^{-2}). \\
&= f'_Y(y_\tau)\frac{1}{n}\sum_{i=1}^{n}\psi_{Q,i}(y_\tau) + R_{f_Y},
\end{aligned}$$

where

$$R_{f_Y} := \left[\hat{f}'_Y(y_\tau) - f'_Y(y_\tau)\right][\hat{y}_\tau - y_\tau] + o_p(n^{-1/2}) + O_p(n^{-1}h^{-2})$$

and the $o_p(n^{-1/2})$ term is the error of the linear asymptotic representation of $\hat{y}_\tau - y_\tau$.

In order to the obtain the order of R_{f_Y}, we use the following results:

$$\begin{aligned}
\hat{f}'_Y(y) &= f'_Y(y) + O_p\left(\frac{1}{\sqrt{nh^3}} + h^2\right), \\
\hat{y}_\tau &= y_\tau + O_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

The rate on the derivative of the density can be found on page 56 of Pagan and Ullah (1999). Therefore,

$$\begin{aligned}
R_{f_Y} &= o_p(n^{-1/2}) + O_p\left(\frac{1}{\sqrt{nh^3}} + h^2\right)O_p\left(\frac{1}{\sqrt{n}}\right) + O_p(n^{-1}h^{-2}). \\
&= o_p(n^{-1/2}) + O_p\left(n^{-1}h^{-3/2}\right) + O_p\left(n^{-1/2}h^2\right) + O_p(n^{-1}h^{-2}). \\
&= o_p(n^{-1/2}) + O_p\left(n^{-1}h^{-3/2}\right) + O_p(n^{-1}h^{-2}).
\end{aligned}$$

because, since by Assumption 2.5, $h \downarrow 0$, so $O_p\left(n^{-1/2}h^2\right) = o_p(n^{-1/2})$. We need to show that $\sqrt{nh}R_{f_Y} = o_p(1)$. We do this term by term. First,

$$\sqrt{nh} \times o_p(n^{-1/2}) = o_p(h^{1/2}) = o_p(1)$$

because $h \downarrow 0$. Second,

$$\sqrt{nh} \times O_p\left(n^{-1}h^{-3/2}\right) = O_p\left(n^{-1/2}h^{-1}\right) = o_p(1)$$

as long as $nh^2 \uparrow \infty$, which is guaranteed by Assumption 2.5, since it is implied by $nh^3 \uparrow \infty$. Finally,

$$\sqrt{nh} \times O_p(n^{-1}h^{-2}) = O_p(n^{-1/2}h^{-3/2}) = o_p(1)$$

since by Assumption 2.5 $nh^3 \uparrow \infty$. Therefore, $\sqrt{nh}R_{f_Y} = o_p(1)$. □

Proof of Lemma 2.4. We have the following decomposition:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \hat{\alpha})}{\partial z} - E\left[\frac{\partial P(Z, X, \alpha_0)}{\partial z}\right] = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \hat{\alpha})}{\partial z} - \frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \alpha_0)}{\partial z} + \frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \alpha_0)}{\partial z} - E\left[\frac{\partial P(Z, X, \alpha_0)}{\partial z}\right].$$

Under the assumption of finite variance for $\frac{\partial P(Z,X,\alpha_0)}{\partial z}$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \alpha_0)}{\partial z} - E\left[\frac{\partial P(Z, X, \alpha_0)}{\partial z}\right] = \frac{1}{n}\sum_{i=1}^{n}\psi_{\partial P,i} = O_p(n^{-1/2}),$$

where

$$\psi_{\partial P,i} = \frac{\partial P(Z_i, X_i, \alpha_0)}{\partial z} - E\left[\frac{\partial P(Z, X, \alpha_0)}{\partial z}\right].$$

For the first term, we have by applying the mean value theorem coordinate-wise

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \hat{\alpha})}{\partial z} - \frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i, X_i, \alpha_0)}{\partial z} = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 P(Z_i, X_i, \tilde{\alpha})}{\partial \alpha \partial z}\right)'(\hat{\alpha} - \alpha_0),$$

where $\tilde{\alpha}$ is a vector with (not necessarily equal) coordinates between α_0 and $\hat{\alpha}$. Under the uniform law of large numbers given in the lemma and the continuity of $\alpha \mapsto E\left[\frac{\partial^2 P(Z,X,\alpha)}{\partial \alpha \partial z}\right]$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 P(Z_i,X_i,\tilde{\alpha})}{\partial\alpha\partial z}\xrightarrow{P}E\left[\frac{\partial^2 P(Z,X,\alpha_0)}{\partial\alpha\partial z}\right]. \tag{A.37}$$

Using (A.37) together with (2.25), we obtain

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i,X_i,\hat{\alpha})}{\partial z}-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i,X_i,\alpha_0)}{\partial z}\\
=&\left\{E\left[\frac{\partial^2 P(Z,X,\alpha_0)}{\partial\alpha\partial z}\right]'+o_p(1)\right\}\left\{\mathbb{P}_n\psi_{\alpha_0}+o_p(n^{-1/2})\right\}\\
=&\ E\left[\frac{\partial^2 P(Z,X,\alpha_0)}{\partial\alpha\partial z}\right]'\mathbb{P}_n\psi_{\alpha_0}+o_p(n^{-1/2}).
\end{aligned}$$

The decomposition is then

$$\begin{aligned}
T_{1n}(\hat{\alpha})-T_1 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial P(Z_i,X_i,\hat{\alpha})}{\partial z}-E\left[\frac{\partial P(Z,X,\alpha_0)}{\partial z}\right]\\
&= E\left[\frac{\partial^2 P(Z,X,\alpha_0)}{\partial\alpha\partial z}\right]'\mathbb{P}_n\psi_{\alpha_0}+\mathbb{P}_n\psi_{\partial P}+o_p(n^{-1/2}).
\end{aligned}$$

□

Proof of Lemma 2.5. Recall that

$$\begin{aligned}
m_0\left(y_\tau,\tilde{w}\left(\alpha_\theta\right)\right) &:= m_0\left(y_\tau,P(w,\alpha_\theta),x\right)\\
&= E\left[\mathbb{1}\left\{Y\le y_\tau\right\}|P(W,\alpha_\theta)=P(w,\alpha_\theta),X=x\right].
\end{aligned}$$

In order to emphasize the dual roles of α_θ, we define

$$\tilde{m}_0(y_\tau,u,x;P(\cdot,\alpha_{\theta_2}))=E\left[\mathbb{1}\left\{Y\le y_\tau\right\}|P(W,\alpha_{\theta_2})=u,X=x\right].$$

Since y_τ is fixed, we regard $\tilde{m}_0$ as a function of (u,x) that depends on the function $P(\cdot,\alpha_{\theta_2})$.

Then

$$\begin{aligned}\tilde{m}_0(y_\tau, P(w,\alpha_{\theta_1}), x; P(\cdot,\alpha_{\theta_2}))|_{\theta_1=\theta_2=\theta} &= E\left[\mathbb{1}\{Y \le y_\tau\} | P(W,\alpha_\theta) = P(w,\alpha_\theta), X = x\right] \\ &= m(y_\tau, P(w,\alpha_\theta), x).\end{aligned}$$

As in Hahn and Ridder (2013), we employ $\tilde{m}_0(y_\tau, u, x; P(\cdot,\alpha_{\theta_2}))$ as an expositional device only.

The functional of interest is

$$\begin{aligned}\mathscr{H}[m_0] &=: \mathscr{H}[m_0(y_\tau, P(\cdot,\alpha_\theta), \cdot)] = \int_{\mathscr{W}} \frac{\partial m_0(y_\tau, P(w,\alpha_\theta), x)}{\partial z} f_W(w)\, dw \\ &= \int_{\mathscr{W}} \left.\frac{\partial \tilde{m}_0(y_\tau, P(w,\alpha_{\theta_1}), x; P(\cdot,\alpha_{\theta_2}))}{\partial z}\right|_{\theta_1=\theta_2=\theta} f_W(w)\, dw.\end{aligned}$$

Under Condition (iii) of the lemma, we can exchange $\frac{\partial}{\partial \alpha_\theta}$ with E and obtain

$$\begin{aligned}\left.\frac{\partial}{\partial \alpha_\theta}\mathscr{H}[m_0]\right|_{\theta=\theta_0} &= \int_{\mathscr{W}} \frac{\partial}{\partial z} \left.\frac{\partial \tilde{m}_0(y_\tau, P(w,\alpha_{\theta_1}), x; P(\cdot,\alpha_{\theta_2}))}{\partial \alpha_{\theta_1}}\right|_{\theta_1=\theta_2=\theta_0} f_W(w)\, dw \\ &+ \int_{\mathscr{W}} \frac{\partial}{\partial z} \left.\frac{\partial \tilde{m}_0(y_\tau, P(w,\alpha_{\theta_1}), x; P(\cdot,\alpha_{\theta_2}))}{\partial \alpha_{\theta_2}}\right|_{\theta_1=\theta_2=\theta_0} f_W(w)\, dw \\ &= \int_{\mathscr{W}} \frac{\partial}{\partial z}\left[\tilde{m}'_{0,\alpha}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right)\right] f_W(w)\, dw,\end{aligned}$$

where

$$\begin{aligned}\tilde{m}'_{0,\alpha}(y_\tau, P(w,\alpha_\theta), x; P(\cdot,\alpha_\theta)) &= \frac{\partial \tilde{m}_0(y_\tau, P(w,\alpha_{\theta_1}), x; P(\cdot,\alpha_\theta))}{\partial \alpha_{\theta_1}} \\ &+ \left.\frac{\partial \tilde{m}_0(y_\tau, P(w,\alpha_\theta), x; P(\cdot,\alpha_{\theta_2}))}{\partial \alpha_{\theta_2}}\right|_{\theta_1=\theta_2=\theta}.\end{aligned}$$

Under Condition (i) of the lemma, we can do integration by parts, and we have

$$
\begin{aligned}
& \int_{\mathscr{W}} \frac{\partial}{\partial z}\left[\tilde{m}_{0,\alpha}^{\prime}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right)\right] f_W(w)\, dw \\
= & \int_{\mathscr{X}} \int_{z_L(x)}^{z_U(x)} \frac{\partial}{\partial z}\left[\tilde{m}_{0,\alpha}^{\prime}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right)\right] f_{Z|X}(z|x)\, dz \cdot f_X(x)\, dx \\
= & \int_{\mathscr{X}} \tilde{m}_{0,\alpha}^{\prime}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right) f_{Z|X}(z|x)\Big|_{z_L(x)}^{z_U(x)} f_X(x)\, dx \\
- & \int_{\mathscr{W}} \tilde{m}_{0,\alpha}^{\prime}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right) \frac{\partial \log f_{Z|X}(z|x)}{\partial z} f_W(w)\, dw \\
= & -\int_{\mathscr{W}} \tilde{m}_{0,\alpha}^{\prime}\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right) \frac{\partial \log f_W(w)}{\partial z} f_W(w)\, dw.
\end{aligned}
$$

Define

$$
\nu\left(u, x; P(\cdot,\alpha_\theta)\right) = E\left[\left.\frac{\partial \log f_W(W)}{\partial z}\right| P(W,\alpha_\theta) = u, X = x\right].
$$

By the law of iterated expectations, we have

$$
\begin{aligned}
& \int_{\mathscr{W}} \tilde{m}_0\left(y_\tau, P(w,\alpha_\theta), x; P\left(\cdot,\alpha_\theta\right)\right) \nu\left(P(w,\alpha_\theta), x; P(\cdot,\alpha_{\theta_0})\right) f_W(w)\, dw \\
= \; & E\left[\mathbb{1}\left\{Y \le y_\tau\right\} \nu\left(P(W,\alpha_\theta), X; P(\cdot,\alpha_{\theta_0})\right)\right].
\end{aligned}
$$

Differentiating the above with respect to α_θ and evaluating the resulting equation at $\theta = \theta_0$, we have

$$
\begin{aligned}
& E\left[\left.\frac{\partial \tilde{m}_0(y_\tau, P(W,\alpha_{\theta_1}), X; P\left(\cdot,\alpha_{\theta_0}\right))}{\partial \alpha_{\theta_1}}\right|_{\theta_1=\theta_0} \nu\left(P(W,\alpha_\theta), X; P(\cdot,\alpha_{\theta_0})\right)\right] \\
+ \; & E\left[\left.\frac{\partial \tilde{m}_0(y_\tau, P(W,\alpha_{\theta_0}), X; P\left(\cdot,\alpha_{\theta_2}\right))}{\partial \alpha_{\theta_2}}\right|_{\theta_2=\theta_0} \nu\left(P(W,\alpha_\theta), X; P(\cdot,\alpha_{\theta_0})\right)\right] \\
= \; & E\left\{\left[\mathbb{1}\left\{Y \le y_\tau\right\} - m\left(y_\tau, P(W,\alpha_{\theta_0}), X\right)\right] \left.\frac{\partial \nu\left(P(W,\alpha_\theta), X; P(\cdot,\alpha_{\theta_0})\right)}{\partial \alpha_\theta}\right|_{\theta=\theta_0}\right\} \qquad \text{(A.38)}
\end{aligned}
$$

where we have used Condition (iii) to exchange the differentiation with the expectation.

Using (A.38) and Condition (ii) of the lemma, we have

$$
\begin{aligned}
& \int_{\mathscr{W}} \frac{\partial}{\partial z}\left[\tilde{m}_{0,\alpha}'\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right)\right] f_W(w)\,dw \\
= \; & -\int_{\mathscr{W}} \tilde{m}_{0,\alpha}'\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right) \frac{\partial \log f_W(w)}{\partial z} f_W(w)\,dw \\
= \; & \int_{\mathscr{W}} \tilde{m}_{0,\alpha}'\left(y_\tau, P(w,\alpha_{\theta_0}), x; P\left(\cdot,\alpha_{\theta_0}\right)\right) \nu\left(P(w,\alpha_{\theta_0}), x; P(\cdot,\alpha_{\theta_0})\right) f_W(w)\,dw \\
= \; & E\left\{\left[\mathbb{1}\{Y \le y_\tau\} - m\left(y_\tau, P(W,\alpha_{\theta_0}), X\right)\right] \frac{\partial \nu\left(P(w,\alpha_{\theta_0}), x; P(\cdot,\alpha_{\theta_0})\right)}{\partial \alpha_{\theta_0}}\right\} \\
= \; & 0.
\end{aligned}
$$

This implies that

$$
\frac{\partial}{\partial \theta} E\left[\frac{\partial m_0(y_\tau, P(Z,X,\alpha_\theta), X)}{\partial z}\right]\Bigg|_{\theta=\theta_0} = 0.
$$

□

Proof of Lemma 2.6. First, we prove that the decomposition in (2.34) is valid. We start by showing that

$$
T_{2,\theta} = E_\theta\left[\frac{\partial m_\theta(y_{\tau,\theta}, \tilde{W}(\alpha_\theta))}{\partial z}\right]
$$

is differentiable at θ_0. For this, it suffices to show that each of the four derivatives below exists at $\theta = \theta_0$:

$$
\begin{aligned}
& \frac{\partial}{\partial \theta} E_\theta\left[\frac{\partial m_0(y_\tau, \tilde{W}(\alpha_0))}{\partial z}\right]; \frac{\partial}{\partial \theta} E\left[\frac{\partial m_\theta(y_\tau, \tilde{W}(\alpha_0))}{\partial z}\right]; \\
& \frac{\partial}{\partial \theta} E\left[\frac{\partial m_0(y_{\tau,\theta}, \tilde{W}(\alpha_0))}{\partial z}\right]; \frac{\partial}{\partial \theta} E\left[\frac{\partial m_0(y_\tau, \tilde{W}(\alpha_\theta))}{\partial z}\right].
\end{aligned}
\tag{A.39}
$$

By Lemma 2.5, the last derivative exists and is equal to zero at $\theta = \theta_0$. We deal with the rest three derivatives in (A.39) one at a time. Consider the first derivative. Under Conditions (i) and (ii) of the lemma, $E_\theta\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial z}\right]$ is differentiable in θ and

$$\frac{\partial}{\partial\theta}E_\theta\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial z}\right]\Bigg|_{\theta=\theta_0}=E\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial z}S(O)\right].$$

Hence, the contribution associated with the first derivative is simply the influence function of $T_{2n}(y_\tau,m_0,\alpha_0)-T_2$.

Now, for the second derivative in (A.39), Theorem 7.2 in Newey (1994) shows that the assumptions of the lemma imply the following:

1. There is a function $\gamma_m(o)$ and a measure $\hat{F}_m$ such that $E[\gamma_m(O)]=0$, $E[\gamma_m(O)^2]<\infty$, and for all $\hat{m}$ such that $\|\hat{m}-m_0\|$ is small enough,

$$E\left[\frac{\partial\hat{m}(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}-\frac{\partial m_0(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]=\int\gamma_m(o)d\hat{F}_m(o).$$

2. The following approximation holds

$$\int\gamma_m(o)d\hat{F}_m(o)=\frac{1}{n}\sum_{i=1}^{n}\gamma_m(O_i)+o_p(n^{-1/2}).$$

For a parametric submodel F_θ, we then have, when θ is close enough to θ_0 :

$$\begin{aligned}&\frac{1}{\theta-\theta_0}E\left[\frac{\partial m_\theta(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}-\frac{\partial m_0(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]\\ =&\frac{1}{\theta-\theta_0}\int\gamma_m(o)d\left[F_\theta(o)-F_{\theta_0}(o)\right],\end{aligned}$$

since $E[\gamma_m(O)]=\int\gamma_m(o)dF_{\theta_0}(o)=0$. If $\int\gamma_m(o)^2dF_\theta(o)$ is bounded in a neighborhood $\theta=\theta_0$, then, by Lemma 7.2 in Ibragimov and Hasminskii (1981), the second derivative exists and satisfies

$$\frac{\partial}{\partial\theta}E\left[\frac{\partial m_\theta(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]\Bigg|_{\theta=\theta_0}=\frac{\partial}{\partial\theta}\int\gamma_m(o)dF_\theta(o)\Bigg|_{\theta=\theta_0}=E[\gamma_m(O)S(O)].$$

This shows that $\gamma_m(o)$ is the influence function of $E\left[\frac{\partial \hat{m}(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]$. That is

$$E\left[\frac{\partial \hat{m}(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]-E\left[\frac{\partial m_0(y_\tau,P(Z,X,\alpha_0),X)}{\partial z}\right]=\frac{1}{n}\sum_{i=1}^{n}\gamma_m(O_i)+o_p(n^{-1/2}).$$

This, combined with the stochastic equicontinuity assumption, implies that

$$T_{2n}(y_\tau,\hat{m},\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0)=\frac{1}{n}\sum_{i=1}^{n}\gamma_m(O_i)+o_p(n^{-1/2}),$$

and $\gamma_m(o)$ is the influence function of $T_{2n}(y_\tau,\hat{m},\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0)$.

Now, the dominating condition in condition (iv) ensures that the third derivative in (A.39) exists and

$$\frac{\partial}{\partial\theta}E\left[\frac{\partial m_0(y_{\tau,\theta},\tilde{W}(\alpha_0))}{\partial z}\right]\bigg|_{\theta=\theta_0}=E\left[\frac{\partial^2 m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial y_\tau\partial z}\right]\frac{\partial y_{\tau,\theta}}{\partial\theta}\bigg|_{\theta=\theta_0}.$$

Given the approximation

$$\hat{y}_\tau-y_\tau=\mathbb{P}_n\psi_Q(y_\tau)+o_p(n^{-1/2}),$$

from Lemma 2.2, we have

$$\frac{\partial y_{\tau,\theta}}{\partial\theta}\bigg|_{\theta=\theta_0}=E\left[\psi_Q(y_\tau)S(O)\right].$$

Hence,

$$\frac{\partial}{\partial\theta}E\left[\frac{\partial m_0(y_{\tau,\theta},\tilde{W}(\alpha_0))}{\partial z}\right]\bigg|_{\theta=\theta_0}=E\left[\frac{\partial^2 m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial y_\tau\partial z}\right]E\left[\psi_Q(y_\tau)S(O)\right].$$

This gives us the contribution from the estimation of y_τ. Alternatively, this expression gives us the influence function of

$$E\left[\frac{\partial m_0(\hat{y}_\tau,\tilde{W}(\alpha_0))}{\partial z}\right]-E\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial z}\right],$$

because

$$E\left[\frac{\partial m_0(\hat{y}_\tau,\tilde{W}(\alpha_0))}{\partial z}\right]-E\left[\frac{\partial m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial z}\right]=E\left[\frac{\partial^2 m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial y_\tau \partial z}\right](\hat{y}_\tau-y_\tau)+o_p(n^{-1/2}).$$

Using the stochastic equicontinuity assumption, we then get that

$$\begin{aligned}
T_{2n}(\hat{y}_\tau,m_0,\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0) &= E\left[\frac{\partial^2 m_0(y_\tau,\tilde{W}(\alpha_0))}{\partial y_\tau \partial z}\right](\hat{y}_\tau-y_\tau)+o_p(n^{-1/2})\\
&= E\left[\frac{\partial f_{Y|\tilde{W}(\alpha_0)}(y_\tau|\tilde{W}(\alpha_0))}{\partial z}\right]\mathbb{P}_n\psi_Q(y_\tau)+o_p(n^{-1/2}).
\end{aligned}$$

To sum up, we have shown that

$$\begin{aligned}
T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})-T_2 &= [T_{2n}(y_\tau,m_0,\alpha_0)-T_2]+[T_{2n}(y_\tau,\hat{m},\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0)]\\
&+ [T_{2n}(\hat{y}_\tau,m_0,\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0)]+o_p(n^{-1/2}).
\end{aligned}$$

To obtain the influence function of the first and second terms in the right-hand side of (2.34), we just need to invoke Theorem 7.2 in Newey (1994) to obtain

$$\begin{aligned}
T_{2n}(y_\tau,\hat{m},\alpha_0)-T_2 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_0(y_\tau,\tilde{W}_i(\alpha_0))}{\partial z}-T_2\\
&- \mathbb{P}_n\left[\mathbb{1}\{Y\le y_\tau\}-m_0(y_\tau,\tilde{W}(\alpha_0))\right]E\left[\frac{\partial \log f_W(W)}{\partial z}\middle|\tilde{W}(\alpha_0)\right]\\
&+ o_p(n^{-1/2}), \qquad\qquad (A.40)
\end{aligned}$$

because

$$T_{2n}(y_\tau,\hat{m},\alpha_0)-T_2=T_{2n}(y_\tau,m_0,\alpha_0)-T_2+T_{2n}(y_\tau,\hat{m},\alpha_0)-T_{2n}(y_\tau,m_0,\alpha_0).$$

The first term in (A.40) is simply the influence function of a sample mean. The second term in (A.40) is the adjustment due to the estimation of m_0. □

Proof of Theorem 2.5. Consider the following difference

$$\begin{aligned}
M_{\tau,z} - \hat{M}_{\tau,z} &= \frac{1}{\hat{f}_Y(\hat{y}_\tau)} \frac{T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha})}{T_{1n}(\hat{\alpha})} - \frac{1}{f_Y(y_\tau)} \frac{T_2}{T_1} \\
&= \frac{T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) f_Y(y_\tau) T_1 - \hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) T_2}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \\
&= \frac{f_Y(y_\tau) T_1}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \left[T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 \right] \\
&- \frac{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) T_2 - f_Y(y_\tau) T_1 T_2}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \\
&= \frac{f_Y(y_\tau) T_1}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \left[T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 \right] \\
&- \frac{T_{1n}(\hat{\alpha}) T_2}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau) \right] \\
&- \frac{T_2 f_Y(y_\tau)}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) f_Y(y_\tau) T_1} \left[T_{1n}(\hat{\alpha}) - T_1 \right]. \qquad \text{(A.41)}
\end{aligned}$$

We can rearrange (A.41) as

$$\begin{aligned}
\hat{M}_{\tau,z} - M_{\tau,z} &= \frac{T_2}{\hat{f}_Y(\hat{y}_\tau) f_Y(y_\tau) T_1} \left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau) \right] \\
&+ \frac{T_2}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha}) T_1} \left[T_{1n}(\hat{\alpha}) - T_1 \right] \\
&- \frac{1}{\hat{f}_Y(\hat{y}_\tau) T_{1n}(\hat{\alpha})} \left[T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 \right]. \qquad \text{(A.42)}
\end{aligned}$$

By appropriately defining the remainders, we can express (A.42) as

$$\begin{aligned}
\hat{M}_{\tau,z} - M_{\tau,z} &= \frac{T_2}{f_Y(y_\tau)^2 T_1} \left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau) \right] \\
&+ \frac{T_2}{f_Y(y_\tau) T_1^2} \left[T_{1n}(\hat{\alpha}) - T_1 \right] \\
&- \frac{1}{f_Y(y_\tau) T_1} \left[T_{2n}(\hat{y}_\tau, \hat{m}, \hat{\alpha}) - T_2 \right] + R_1 + R_2 + R_3. \qquad \text{(A.43)}
\end{aligned}$$

The definitions of R_1, R_2 and R_3 can be found below. Now we are ready to separate the contribution of each stage of the estimation. We shall use equations (2.22), (2.26), and (2.34).

$$\begin{aligned}
\hat{M}_{\tau,z} - M_{\tau,z} &= \frac{T_2}{f_Y(y_\tau)^2 T_1}\left[\hat{f}_Y(y_\tau) - f_Y(y_\tau)\right] + \frac{T_2}{f_Y(y_\tau)^2 T_1}\left[f_Y(\hat{y}_\tau) - f_Y(y_\tau)\right] \\
&+ \frac{T_2}{f_Y(y_\tau)T_1^2}\left[T_{1n}(\alpha_0) - T_1\right] + \frac{T_2}{f_Y(y_\tau)T_1^2}\left\{E\left[\frac{\partial P(W,\hat{\alpha})}{\partial z}\right] - T_1\right\} \\
&- \frac{1}{f_Y(y_\tau)T_1}\left[T_{2n}(y_\tau, m_0, \alpha_0) - T_2\right] \\
&- \frac{1}{f_Y(y_\tau)T_1}\left[T_{2n}(y_\tau, \hat{m}, \alpha_0) - T_{2n}(y_\tau, m_0, \alpha_0)\right] \\
&- \frac{1}{f_Y(y_\tau)T_1}\left[T_{2n}(\hat{y}_\tau, m_0, \alpha_0) - T_{2n}(y_\tau, m_0, \alpha_0)\right] \\
&+ R_1 + R_2 + R_3 + \frac{T_2}{f_Y(y_\tau)^2 T_1} R_{f_Y} + o_p(n^{-1/2}). \qquad \text{(A.44)}
\end{aligned}$$

Finally, we establish the rate for the remainders R_1, R_2 and R_3 in (A.43). We deal with each component of the remainder separately. The first remainder is

$$\begin{aligned}
R_1 &= \frac{T_2}{\hat{f}_Y(\hat{y}_\tau) f_Y(y_\tau) T_1}\left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)\right] - \frac{T_2}{f_Y(y_\tau)^2 T_1}\left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)\right] \\
&= \frac{T_2}{f_Y(y)T_1}\left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)\right]\left[\frac{1}{\hat{f}_Y(\hat{y}_\tau)} - \frac{1}{f_Y(y_\tau)}\right] \\
&= -\frac{T_2}{f_Y(y)^2 T_1 \hat{f}_Y(\hat{y}_\tau)}\left[\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)\right]^2 \\
&= O_p\left(|\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau)|^2\right). \qquad \text{(A.45)}
\end{aligned}$$

The second remainder is

$$\begin{aligned}
R_2 &= \frac{T_2}{\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})T_1}\left[T_{1n}(\hat{\alpha})-T_1\right]-\frac{T_2}{f_Y(y_\tau)T_1^2}\left[T_{1n}(\hat{\alpha})-T_1\right] \\
&= \frac{T_2}{T_1}\left[T_{1n}(\hat{\alpha})-T_1\right]\left[\frac{1}{\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})}-\frac{1}{f_Y(y_\tau)T_1}\right] \\
&= \frac{T_2}{T_1}\left[T_{1n}(\hat{\alpha})-T_1\right]\left[\frac{f_Y(y_\tau)T_1-\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})}{\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})f_Y(y_\tau)T_1}\right] \\
&= \frac{T_2}{T_1}\left[T_{1n}(\hat{\alpha})-T_1\right]\left[\frac{f_Y(y_\tau)(T_1-T_{1n}(\hat{\alpha}))-T_{1n}(\hat{\alpha})(\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau))}{\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})f_Y(y_\tau)T_1}\right] \\
&= O_p\left(|T_{1n}(\hat{\alpha})-T_1|^2\right)+O_p\left(|T_{1n}(\hat{\alpha})-T_1||\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right) \\
&= O_p\left(n^{-1}\right)+O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right). \qquad \text{(A.46)}
\end{aligned}$$

The third remainder is

$$\begin{aligned}
R_3 &= \frac{1}{f_Y(y_\tau)T_1}\left[T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})-T_2\right]-\frac{1}{\hat{f}_Y(\hat{y}_\tau)T_{1n}(\hat{\alpha})}\left[T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})-T_2\right] \\
&= O_p\left(|T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})-T_2||T_{1n}(\hat{\alpha})-T_1|\right) \\
&+ O_p\left(|T_{2n}(\hat{y}_\tau,\hat{m},\hat{\alpha})-T_2||\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right) \\
&= O_p\left(n^{-1}\right)+O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right) \qquad \text{(A.47)}
\end{aligned}$$

because it has the same denominator as R_2 in (A.46). Finally, we compute the rate for $\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)$. To do so, we use the results in Lemma 2.3. Equation (2.22) tells us

$$\begin{aligned}
\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau) &= \hat{f}_Y(y_\tau)-f_Y(y_\tau)+f_Y(\hat{y}_\tau)-f_Y(y_\tau)+R_{f_Y} \\
&= O_p(n^{-1/2}h^{-1/2})+O(h^2)+o_p(h^2)+O_p(n^{-1/2})+O_p(|R_{f_Y}|) \\
&= O_p(n^{-1/2}h^{-1/2})+O(h^2)+o_p(h^2)+O_p(n^{-1/2})+O_p(|R_{f_Y}|) \\
&= O_p(n^{-1/2}h^{-1/2})+O(h^2)+o_p(n^{-1/2})+O_p(|R_{f_Y}|). \qquad \text{(A.48)}
\end{aligned}$$

Thus we have

$$O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|^2\right)=O_p(n^{-1}h^{-1})+O(h^4)+o_p(n^{-1})+O_p(|R_{f_Y}|^2).$$

The remainder R_M is defined as

$$R_M:=R_1+R_2+R_3+\frac{T_2}{f_Y(y_\tau)^2T_1}R_{f_Y}+o_p(n^{-1/2})+o_p(h^2).$$

So,

$$\begin{aligned} R_M &= O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|^2\right)+O_p\left(n^{-1}\right)+O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right)\\ &+ O_p\left(|R_{f_Y}|\right)+o_p(n^{-1/2})+o_p(h^2)\\ &= O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|^2\right)+O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right)\\ &+ O_p\left(|R_{f_Y}|\right)+o_p(n^{-1/2})+o_p(h^2), \end{aligned} \tag{A.49}$$

because $O_p(n^{-1})$ is $o_p(n^{-1/2})$ as $n\to\infty$.

Now we show that $\sqrt{nh}R_M=o_p(1)$ under Assumption 2.5. We do this term by term in (A.49):

$$\sqrt{nh}o_p(h^2)=o_p(n^{1/2}h^{5/2})=o_p(1) \text{ as } nh^5\downarrow 0;$$

$$\sqrt{nh}o_p(n^{-1/2})=o_p(h^{1/2})=o_p(1) \text{ as } h\downarrow 0;$$

$$\sqrt{nh}O_p\left(|R_{f_Y}|\right)=o_p(1) \text{ by Lemma 2.3;}$$

$$\sqrt{nh}O_p\left(n^{-1/2}|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right)=h^{1/2}O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|\right)=o_p(1),$$

since $\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)=o_p(1)$. Finally,

$$\sqrt{nh}O_p\left(|\hat{f}_Y(\hat{y}_\tau)-f_Y(y_\tau)|^2\right)=o_p(1),$$

since $\hat{f}_Y(\hat{y}_\tau) - f_Y(y_\tau) = O_p(n^{-1/2}h^{-1/2} + h^2 + n^{-1/2} + R_{f_Y})$ by (2.22). □

Details of Example 1.8. The propensity score under the new policy regime is

$$P^{\mu}_{\delta}(w) = \Pr\left[V \leq \mu(w) + s_\mu(\delta)\,|W = w\right] = F_{V|W}(\mu(w) + s_\mu(\delta)\,|w),$$

and so

$$\left.\frac{\partial P^{\mu}_{\delta}(w)}{\partial \delta}\right|_{\delta=0} = \frac{f_V(\mu(w))}{E(f_V(\mu(W)))}.$$

By Theorem 1.3, we have

$$\begin{aligned} M_{\tau,\mu} &= \frac{1}{E(f_V(\mu(W)))f_Y(y_\tau)} \int_{\mathscr{W}} E\left[\mathbb{1}\{Y(0) \leq y_\tau\}\,|V = \mu(w), W = w\right] f_V(\mu(w)) f_W(w) dw \\ &- \frac{1}{E(f_V(\mu(W)))f_Y(y_\tau)} \int_{\mathscr{W}} E\left[\mathbb{1}\{Y(1) \leq y_\tau\}\,|V = \mu(w), W = w\right] f_V(\mu(w)) f_W(w) dw. \end{aligned}$$

Using the potential outcome equations, we get

$$\begin{aligned} E\left[\mathbb{1}\{Y(0) \leq y_\tau\}\,|V = \mu(w), W = w\right] &= E\left[\mathbb{1}\{q(w) + U_0 \leq y_\tau\}\,|V = \mu(w)\right] \\ &= \Pr(U_0 \leq y_\tau - q(w)|V = \mu(w)) \\ &= F_{U_0|V}(y_\tau - q(w)|\mu(w)) \end{aligned}$$

and

$$\begin{aligned} E\left[\mathbb{1}\{Y(1) \leq y_\tau\}\,|V = \mu(w), W = w\right] &= E\left[\mathbb{1}\{q(w) + \beta + U_1 \leq y_\tau\}\,|V = \mu(w)\right] \\ &= \Pr(U_1 \leq y_\tau - q(w) - \beta|V = \mu(w)) \\ &= F_{U_1|V}(y_\tau - q(w) - \beta|\mu(w)). \end{aligned}$$

Hence,

$$\begin{aligned}
M_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} F_{U_0|V}(y_\tau - q(w)|\mu(w)))\frac{f_V(\mu(w))f_W(w)}{\int_{\mathscr{W}} f_V(\mu(w))f_W(w)dw}dw \\
&- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \left\{F_{U_1|V}(y_\tau - q(w) - \beta|\mu(w))\right\}\frac{f_V(\mu(w))f_W(w)}{\int_{\mathscr{W}} f_V(\mu(w))f_W(w)dw}dw \\
&= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \left[\int_{-\infty}^{y_\tau - q(w)} f_{U_0|V}(u|\mu(w))du\right]\tilde{f}_W(w)dw \\
&- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \left[\int_{-\infty}^{y_\tau - q(w) - \beta} f_{U_1|V}(u|\mu(w))du\right]\tilde{f}_W(w)dw.
\end{aligned}$$

It follows from Corollary 1.4 that the apparent effect is

$$\begin{aligned}
A_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{-\infty}^{\infty} E\left[\mathbb{1}\left\{Y \le y_\tau\right\}|D = 0, W = w\right] f_W(w)dw \\
&- \frac{1}{f_Y(y_\tau)}\int_{-\infty}^{\infty} E\left[\mathbb{1}\left\{Y \le y_\tau\right\}|D = 1, W = w\right] f_W(w)dw.
\end{aligned}$$

Using the explicit forms of the potential outcomes, we get

$$\begin{aligned}
E\left[\mathbb{1}\left\{Y \le y_\tau\right\}|D = 0, W = w\right] &= E\left[\mathbb{1}\left\{Y(0) \le y_\tau\right\}|D = 0, W = w\right] \\
&= E\left[\mathbb{1}\left\{q(W) + U_0 \le y_\tau\right\}|D = 0, W = w\right] \\
&= E\left[\mathbb{1}\left\{q(W) + U_0 \le y_\tau\right\}|V > \mu(w), W = w\right] \\
&= \Pr(U_0 \le y_\tau - q(w)|V > \mu(w)) \\
&= \frac{\int_{\mu(w)}^{\infty}\left[\int_{-\infty}^{y_\tau - q(w)} f_{U_0|V}(u|v)du\right] f_V(v)dv)}{1 - F_V(\mu(w))}
\end{aligned}$$

and

$$\begin{aligned}
E\left[\mathbb{1}\left\{Y \leq y_\tau\right\}|D=1, W=w\right] &= E\left[\mathbb{1}\left\{Y(1) \leq y_\tau\right\}|D=1, W=w\right] \\
&= E\left[\mathbb{1}\left\{q(W)+\beta+U_1 \leq y_\tau\right\}|D=1, W=w\right] \\
&= E\left[\mathbb{1}\left\{q(W)+\beta+U_1 \leq y_\tau\right\}|V \leq \mu(W), W=w\right] \\
&= \Pr(U_1 \leq y_\tau - q(w) - \beta | V \leq \mu(w)) \\
&= \frac{\int_{-\infty}^{\mu(w)}\left[\int_{-\infty}^{y_\tau-q(w)-\beta} f_{U_1|V}(u|v)du\right] f_V(v)\,dv}{F_V(\mu(w))}.
\end{aligned}$$

Hence

$$\begin{aligned}
A_{\tau,\mu} &= \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \frac{\int_{\mu(w)}^{\infty}\left[\int_{-\infty}^{y_\tau-q(w)} f_{U_0|V}(u|v)du\right] f_V(v)\,dv}{1-F_V(\mu(w))} f_W(w)dw \\
&- \frac{1}{f_Y(y_\tau)}\int_{\mathscr{W}} \frac{\int_{-\infty}^{\mu(w)}\left[\int_{-\infty}^{y_\tau-q(w)-\beta} f_{U_1|V}(u|v)du\right] f_V(v)\,dv}{F_V(\mu(w))} f_W(w)dw.
\end{aligned}$$

Now, for the scaling factor $f_Y(\cdot)$, we have a mixture

$$f_Y(y_\tau) = f_{Y(1)|D}(y_\tau|1)\Pr(D=1) + f_{Y(0)|D}(y_\tau|0)\Pr(D=0).$$

The mixing weights are

$$\Pr(D=1) = \Pr(V \leq \mu(W)),$$

and $\Pr(D=1) = 1-\Pr(D=0)$. To obtain the mixing densities, we note that

$$\begin{aligned} F_{Y(0)|D}(y_\tau|0) &= \Pr(Y(0) \leq y_\tau | D=0) = \frac{\Pr(Y(0) \leq y_\tau, D=0)}{\Pr(D=0)} \\ &= \frac{1}{\Pr(D=0)} \Pr(q(W)+U_0 \leq y_\tau, V > \mu(W)) \\ &= \frac{1}{\Pr(D=0)} \int_{\mathscr{W}} \Pr(q(w)+U_0 \leq y_\tau, V > \mu(w)) f_W(w) dw \\ &= \frac{1}{\Pr(D=0)} \int_{\mathscr{W}} \left[F_{U_0}(y_\tau - q(w)) - F_{U_0,V}(y_\tau - q(w), \mu(w))\right] f_W(w) dw. \end{aligned}$$

Hence, the density $f_{Y(0)|D}(y_\tau|0)$ is

$$f_{Y(0)|D}(y_\tau|0) = \frac{1}{\Pr(D=0)} \int_{\mathscr{W}} \left[f_{U_0}(y_\tau - q(w)) - \int_{-\infty}^{\mu(w)} f_{U_0,V}(y_\tau - q(w), \check{w}) d\check{w} \right] f_W(w) dw.$$

For the other case, we have

$$\begin{aligned} F_{Y(1)|D}(y_\tau|1) &= \Pr(Y(1) \leq y_\tau | D=1) = \frac{\Pr(Y(1) \leq y_\tau, D=1)}{\Pr(D=1)} \\ &= \frac{1}{\Pr(D=1)} \Pr(q(W)+\beta+U_1 \leq y_\tau, V \leq \mu(W)) \\ &= \frac{1}{\Pr(D=1)} \int_{\mathscr{W}} \Pr(q(w)+\beta+U_1 \leq y_\tau, V \leq \mu(w)) f_W(w) dw \\ &= \frac{1}{\Pr(D=1)} \int_{\mathscr{W}} F_{U_1,V}(y_\tau - q(w) - \beta, \mu(w)) f_W(w) dw, \end{aligned}$$

and so the density $f_{Y(1)|D}(y_\tau|1)$ is

$$f_{Y(1)|D}(y_\tau|1) = \frac{1}{\Pr(D=1)} \int_{\mathscr{W}} \left[\int_{-\infty}^{\mu(w)} f_{U_1,V}(y_\tau - q(w) - \beta, \check{w}) d\check{w} \right] f_W(w) dw.$$

The density $f_Y(y_\tau)$ is then

$$\begin{aligned} f_Y(y_\tau) &= \int_{\mathscr{W}} \left[f_{U_0}(y_\tau - q(w)) - \int_{-\infty}^{\mu(w)} f_{U_0,V}(y_\tau - q(w), \check{w}) d\check{w} \right] f_W(w) dw \\ &+ \int_{\mathscr{W}} \left[\int_{-\infty}^{\mu(w)} f_{U_1,V}(y_\tau - q(w) - \beta, \check{w}) d\check{w} \right] f_W(w) dw. \end{aligned}$$

$\square$

Proof of Proposition 2.1. Note that for any bounded function $\mathscr{G}(\cdot)$, we have

$$\begin{aligned}
E[\mathscr{G}(Y)|P(W)=P(w),X=x] &= E[\mathscr{G}(Y(1))|D=1,P(W)=P(w),X=x] \\
&\quad \times \Pr(D=1|P(W)=P(w),X=x) \\
&\quad + E[\mathscr{G}(Y(0))|D=0,P(W)=P(w),X=x] \\
&\quad \times \Pr(D=0|P(W)=P(w),X=x).
\end{aligned}$$

But, using $D=\mathbb{1}\{U_D \le P(W)\}$, we have

$$\begin{aligned}
\Pr(D=1|P(W)=P(w),X=x) &= \Pr(U_D \le P(W)|P(W)=P(w),X=x) \\
&= \Pr(U_D \le P(w)|P(W)=P(w),X=x) \\
&= P(w),
\end{aligned}$$

because U_D is independent of W. So,

$$\begin{aligned}
E[\mathscr{G}(Y)|P(W)=P(w),X=x] &= E[\mathscr{G}(Y(1))|D=1,P(W)=P(w),X=x]\,P(w) \\
&\quad + E[\mathscr{G}(Y(0))|D=0,P(W)=P(w),X=x]\,(1-P(w)). \\
&= E[\mathscr{G}(Y(1)))|U_D \le P(w),P(W)=P(w),X=x]\,P(w) \\
&\quad + E[\mathscr{G}(Y(0))|U_D > P(w),P(W)=P(w),X=x]\,(1-P(w)) \\
&= E[\mathscr{G}(Y(1)))|U_D \le P(w),X=x]\,P(w) \\
&\quad + E[\mathscr{G}(Y(0))|U_D > P(w),X=x]\,(1-P(w)),
\end{aligned}$$

where the last line follows because $U=(U_0,U_1)$ is independent of Z given X and U_D (see (2.1)).

Now

$$
\begin{aligned}
E\left[\mathscr{G}(Y(1))|U_D \leq P(w), X=x\right] &= E\left\{E\left[\mathscr{G}(Y(1))|U_D, X=x\right]|U_D \leq P(w), X=x\right\} \\
&= E\left\{E\left[\mathscr{G}(Y(1))|U_D, X=x\right]|U_D \leq P(w)\right\} \\
&= \frac{\int_0^{P(w)} E\left[\mathscr{G}(Y(1))|U_D=u, X=x\right] du}{P(w)},
\end{aligned}
$$

where the first equality uses the law of iterated expectations, the second equality uses the independence of U_D from X, and the last equality uses $U_D \sim$ uniform on $[0,1]$. Similarly,

$$
E\left[\mathscr{G}(Y(0))|U_D > P(w), X=x\right] = \frac{\int_{P(w)}^1 E\left[\mathscr{G}(Y(1))|U_D=u, X=x\right] du}{1-P(w)}.
$$

So we have

$$
\begin{aligned}
E\left[\mathscr{G}(Y)|P(W)=P(w), X=x\right] &= \int_0^{P(w)} E\left[\mathscr{G}(Y(1))|U_D=u, X=x\right] du \\
&+ \int_{P(w)}^1 E\left[\mathscr{G}(Y(0))|U_D=u, X=x\right] du.
\end{aligned}
$$

By taking $\mathscr{G}(\cdot) = \mathbb{1}\left\{\cdot \leq y_\tau\right\}$, we have

$$
\begin{aligned}
E\left[\mathbb{1}\left\{Y \leq y_\tau\right\}|P(W)=P(w), X=x\right] &= \int_0^{P(w)} E\left[\mathbb{1}\left\{Y(1) \leq y_\tau\right\}|U_D=u, X=x\right] du \\
&+ \int_{P(w)}^1 E\left[\mathbb{1}\left\{Y(0) \leq y_\tau\right\}|U_D=u, X=x\right] du.
\end{aligned}
$$

Under Assumptions 1.5(a) and 1.5(b), we can invoke the fundamental theorem of calculus to obtain

$$
\begin{aligned}
\frac{\partial E\left[\mathbb{1}\left\{Y \leq y_\tau\right\}|P(W)=P(w), X=x\right]}{\partial P(w)} &= E\left[\mathbb{1}\left\{Y(1) \leq y_\tau\right\}|U_D=P(w), X=x\right] \\
&- E\left[\mathbb{1}\left\{Y(0) \leq y_\tau\right\}|U_D=P(w), X=x\right] \\
&= \mathrm{MTE}_\tau(P(w), x). \qquad \text{(A.50)}
\end{aligned}
$$

That is,

$$\text{MTE}_\tau(u,x) = \frac{\partial E\left[\mathbb{1}\left\{Y \le y_\tau\right\}|P(W)=u, X=x\right]}{\partial u}$$

for any u such that there is a $w \in \mathscr{W}$ satisfying $P(w) = u$. □

Proof of Equation (2.16). Note that

$$E(Y_\delta) = \int_0^1 E(Y_\delta|\mathscr{P}_\delta = t) f_{\mathscr{P}_\delta}(t)dt. \tag{A.51}$$

Now, consider $E(Y_\delta|\mathscr{P}_\delta = t)$. Depending on the value of $\mathscr{P}_\delta$ relative to the index U_D, we observe the potential outcome $Y(0)$ or $Y(1)$. By the independence of $\mathscr{P}_\delta$ from U_D and the law of iterated expectations, we have

$$\begin{aligned} E(Y_\delta|\mathscr{P}_\delta = t) &= \int_0^1 E(Y_\delta|\mathscr{P}_\delta = t, U_D = u)du \\ &= \int_0^t E(Y(1)|U_D = u)du + \int_t^1 E(Y(0)|U_D = u)du \\ &= \int_0^1 \Big[\mathbb{1}\left\{0 \le u \le t\right\} E(Y(1)|U_D = u) \\ &+ \mathbb{1}\left\{t \le u \le 1\right\} E(Y(0)|U_D = u)\Big] du \end{aligned} \tag{A.52}$$

Plugging (A.52) back into (A.51) we get

$$\begin{aligned} & E(Y_\delta) \\ = & \int_0^1 \left\{ \int_0^1 \Big[\mathbb{1}\left\{0 \le u \le t\right\} E(Y(1)|U_D = u) + \mathbb{1}\left\{t \le u \le 1\right\} E(Y(0)|U_D = u)\Big] du \right\} \\ \times & f_{\mathscr{P}_\delta}(t)dt \\ = & \int_0^1 \left\{ \int_0^1 \Big[\mathbb{1}\left\{u \le t \le 1\right\} E(Y(1)|U_D = u) + \mathbb{1}\left\{0 \le t \le u\right\} E(Y(0)|U_D = u)\Big] f_{\mathscr{P}_\delta}(t)dt \right\} \\ \times & du \\ = & \int_0^1 \Big[(1 - F_{\mathscr{P}_\delta}(u))E(Y(1)|U_D = u) + F_{\mathscr{P}_\delta}(u)E(Y(0)|U_D = u)\Big] du. \end{aligned} \tag{A.53}$$

Going back to (2.15) using (A.53), we get

$$\begin{aligned} \text{PRTE} &= \frac{1}{\delta}\int_0^1 \Big[(1-F_{\mathscr{P}_\delta}(u))E(Y(1)|U_D=u)+F_{\mathscr{P}_\delta}(u)E(Y(0)|U_D=u)\Big]du \\ &- \frac{1}{\delta}\int_0^1 \Big[(1-F_{\mathscr{P}}(u))E(Y(1)|U_D=u)-F_{\mathscr{P}}(u)E(Y(0)|U_D=u)\Big]du \\ &= \frac{1}{\delta}\int_0^1 \text{MTE}(u)(F_{\mathscr{P}}(u)-F_{\mathscr{P}_\delta}(u))du. \end{aligned} \tag{A.54}$$

Taking the limit in (A.54) as $\delta \to 0$ yields

$$\begin{aligned} \text{MPRTE} &= \lim_{\delta\to 0} -\frac{1}{\delta}\int_0^1 \text{MTE}(u)(F_{\mathscr{P}_\delta}(u)-F_{\mathscr{P}}(u))du \\ &= -\int_0^1 \text{MTE}(u)\frac{\partial F_{\mathscr{P}_\delta}(u)}{\partial\delta}\bigg|_{\delta=0} du. \end{aligned}$$

□

Proof of the Equivalence of MPRTE to M_ρ for the Mean Functional. The proof is heuristic, and we do not claim to deal rigorously with all the mathematical issues that arise, even though the proof can be made rigorous with additional technical details. It is hoped that the formal proof given here, combined with the rigorous proof for a special case given in the main text, is enough to convince a reader that the representation we establish here coincides with the existing presentation of the MPRTE.

Note that

$$\begin{aligned} F_{\mathscr{P}_\delta}(u) &= \Pr(\mathscr{P}_\delta \le u) = \Pr(F_{V|X}(\mu(Z+s_z(\delta),X)|X) \le u) \\ &= \int_{\mathscr{W}} \mathbb{1}\left\{F_{V|X}(\mu(z+s_z(\delta),x)|x) \le u\right\} F_W(w)\,dw. \end{aligned}$$

Let G_σ be a smooth CDF such that as $\sigma \to 0$, $G_\sigma(u)$ approaches the step function $\mathbb{1}\{u \ge 0\}$ and its derivative G'_σ approaches the Dirac Delta function G'_0. We have

$$\begin{aligned}
& \lim_{\delta\to 0}\frac{1}{\delta}\left[F_{\mathscr{P}_\delta}(u)-F_{\mathscr{P}}(u)\right] \\
= & \lim_{\delta\to 0}\int_{\mathscr{W}}\frac{\mathbb{1}\left\{F_{V|X}(\mu(z,x)|x)-u>0\right\}-\mathbb{1}\left\{F_{V|X}(\mu(z+s_z(\delta),x)|x)-u>0\right\}}{\delta}f_W(w)\,dw \\
= & \lim_{\delta\to 0}\int_{\mathscr{W}}\lim_{\sigma\to 0}\frac{G_\sigma\left[F_{V|X}(\mu(z,x)|x)-u\right]-G_\sigma\left[F_{V|X}(\mu(z+s_z(\delta),x)|x)-u\right]}{\delta}f_W(w)\,dw \\
= & -\int_{\mathscr{W}}G'_0\left(F_{V|X}(\mu(z,x)|x)-u\right)f_{V|X}\left(\mu(z,x)|x\right)\mu_z(z,x)\left.\frac{\partial s_z(\delta)}{\partial\delta}\right|_{\delta=0}f_W(w)\,dw.
\end{aligned}$$

Therefore,

$$\begin{aligned}
- & \int_0^1\text{MTE}(u,x)\lim_{\delta\to 0}\frac{1}{\delta}\left[F_{\mathscr{P}_\delta}(u)-F_{\mathscr{P}}(u)\right]du \\
= & \int_0^1\int_{\mathscr{W}}\text{MTE}(u,x)G'_0\left(F_{V|X}(\mu(z,x)|x)-u\right)f_{V|X}\left(\mu(z,x)|x\right)\mu_z(z,x)\left.\frac{\partial s_z(\delta)}{\partial\delta}\right|_{\delta=0} \\
\times & f_W(w)\,dwdu \\
= & \int_{\mathscr{W}}\text{MTE}(P(w),x)f_{V|X}\left(\mu(z,x)|x\right)\mu_z(z,x)\left.\frac{\partial s_z(\delta)}{\partial\delta}\right|_{\delta=0}f_W(w)\,dwdu \\
= & \int_{\mathscr{W}}\text{MTE}(P(w),x)\dot{P}(w)f_W(w)\,dwdu \\
= & E\left\{\left[Y(1)-Y(0)\,|U_D=P(W),X\right]\dot{P}(W)\right\}=M_\rho.
\end{aligned}$$

□

Proof of Theorem 3.1. Using the fact that $Y_{D_\delta}=D_\delta Y(1)+(1-D_\delta)Y(0)$, we have

$$\begin{aligned}
F_{Y_{D_\delta}}(y)=&\Pr(D=0,D_\delta=0)F_{Y(0)|D=0,D_\delta=0}(y)+\Pr(D=0,D_\delta=1)F_{Y(1)|D=0,D_\delta=1}(y) \\
&+\Pr(D=1,D_\delta=0)F_{Y(0)|D=1,D_\delta=0}(y)+\Pr(D=1,D_\delta=1)F_{Y(1)|D=1,D_\delta=1}(y).
\end{aligned}$$

Under Assumption 3.1, the probability weights are

$$\Pr(D=0,D_\delta=0)=1-p-\delta,$$
$$\Pr(D=0,D_\delta=1)=\delta,$$
$$\Pr(D=1,D_\delta=0)=0,$$
$$\Pr(D=1,D_\delta=1)=p.$$

Therefore, we rewrite $F_{Y_{D_\delta}}(y)$ as

$$F_{Y_{D_\delta}}(y)=(1-p-\delta)F_{Y(0)|D=0,D_\delta=0}(y)+\delta F_{Y(1)|D=0,D_\delta=1}(y)$$
$$+pF_{Y(1)|D=1,D_\delta=1}(y),$$

We add and subtract $\delta F_{Y(1)|D=1,D_\delta=1}(y)$, to get

$$F_{Y_{D_\delta}}(y)=F_a(y)+\delta\left[F_{Y(1)|D=0,D_\delta=1}(y)-F_{Y(1)|D=1,D_\delta=1}(y)\right].$$

□

Proof of Remark 3.5. We want to show that

$$F_a^{-1}(\tau-\delta)-F_Y^{-1}(\tau)\leq 0,$$

and

$$F_a^{-1}(\tau+\delta)-F_Y^{-1}(\tau)\geq 0.$$

Manipulating equation (3.4) in Theorem 3.1 we can obtain that F_a and F_Y are related by[4]

$$F_a(y) = F_Y(y) + \delta\left[F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1,D_\delta=1}(y)\right].$$

Since $-1 \leq F_{Y(1)|D=0,D_\delta=1}(y) - F_{Y(1)|D=1,D_\delta=1}(y) < 1$, then

$$F_Y(y) - \delta \leq F_a(y) \leq F_Y(y) + \delta.$$

Therefore, we have

$$F_a^{-1}(F_Y(y) - \delta) \leq y \leq F_a^{-1}(F_Y(y) + \delta).$$

Since the previous display is valid for any $y \in \mathscr{Y}$, we set $y = F_Y^{-1}(\tau)$. This implies that

$$F_a^{-1}(\tau - \delta) \leq F_Y^{-1}(\tau) \leq F_a^{-1}(\tau + \delta).$$

Thus, we have that $F_a^{-1}(\tau - \delta) - F_Y^{-1}(\tau) \leq 0$ and $F_a^{-1}(\tau + \delta) - F_Y^{-1}(\tau) \geq 0$. □

Proof of Theorem 3.3. The lower bound is the limit when δ goes to 0 of

$$\frac{F_a^{-1}(\tau - \delta U) - F_Y^{-1}(\tau)}{\delta}. \tag{A.55}$$

We will show that this limit exists and compute its value. Recall that by (3.1) we can simplify the apparent distribution in (3.4) to

$$F_a(y) = (1 - p - \delta)F_{Y|D=0}(y) + (p + \delta)F_{Y|D=1}(y)$$

[4]This follows from noting that $F_Y(y) = (1-p-\delta)F_{Y|D=0,D_\delta=0}(y) + \delta F_{Y|D=0,D_\delta=1}(y) + pF_{Y|D=1,D_\delta=1}(y)$, while by (3.4), $F_a(y) = (1-p-\delta)F_{Y|D=0,D_\delta=0}(y) + (p+\delta)F_{Y|D=1,D_\delta=1}(y)$. So, if we add and subtract $\delta F_{Y|D=0,D_\delta=1}(y)$, we obtain $F_a(y) = F(y) + \delta\left[F_{Y|D=1,D_\delta=1}(y) - F_{Y|D=0,D_\delta=1}(y)\right]$.

We will write $F_{a,\delta}(y)$ to make explicit the fact that the apparent distribution depends on δ. Define

$$g(\delta_1, \delta_2) = F^{-1}_{a,\delta_1}(\tau - \delta_2 U)$$

to emphasize the double role played by δ. The map $\delta_1 \mapsto g(\delta_1, \delta_2)$ for a fixed δ_2 is the composition

$$\delta_1 \in \mathbb{R} \overset{h}{\mapsto} F_{a,\delta_1} \in D[-\infty, \infty] \overset{\Gamma}{\mapsto} F^{-1}_{a,\delta_1}(\tau - \delta_2 U) \in \mathbb{R}.$$

For $[a,b] \subset [-\infty, \infty]$, $D[a,b]$ is the set of all real-valued cadlag functions: right continuous with left limits everywhere in $[a,b]$. $D[a,b]$ is equipped with the uniform norm $\|\cdot\|_\infty$. The first map $h : \delta_1 \in \mathbb{R} \mapsto F_{a,\delta_1} \in D[-\infty, \infty]$ has Hadamard derivative given by $F_{Y|D=1}(y) - F_{Y|D=0}(y)$, while the second map has Hadamard derivative given by (See Lemma 21.3 in van der Vaart (1998))

$$\Gamma'_{F_{a,\delta_1}}[G] = -\frac{G(F^{-1}_{a,\delta_1}(\tau - \delta_2 C))}{f_{a,\delta_1}(F^{-1}_{a,\delta_1}(\tau - \delta_2 U))}.$$

for $G \in D[-\infty, \infty]$ continuous at $F^{-1}_{a,\delta_1}(\tau - \delta_2 C)$. Then, the derivative of the composite map $\delta_1 \mapsto \Gamma \circ h(\delta_1)$ is $\Gamma'_{F_{a,\delta_1}}[h'(\delta_1)]$, which is for a $\delta_2 = 0$

$$\frac{\partial F^{-1}_{a,\delta_1}(\tau)}{\partial \delta_1} = -\frac{F_{Y|D=1}(F^{-1}_{a,\delta_1}(\tau)) - F_{Y|D=0}(F^{-1}_{a,\delta_1}(\tau))}{f_{a,\delta_1}(F^{-1}_{a,\delta_1}(\tau))}.$$

which is continuous at $\delta_1 = 0$. The derivative of the second map $\delta_2 \mapsto g(\delta_1, \delta_2)$, for a fixed $\delta_1 = 0$, can be obtained via the identity

$$F_{a,\delta_1}\left(F^{-1}_{a,\delta_1}(\tau - \delta_2 U)\right) = \tau - \delta_2 U.$$

Differentiating through with respect to δ_2, we obtain

$$\frac{\partial F_a^{-1}(\tau - \delta_2 U)}{\partial \delta_2} = -\frac{U}{f_Y(F_Y^{-1}(\tau - \delta_2 U))}.$$

which is continuous with respect to δ_2.

Therefore, both partial derivatives of the map $(\delta_1, \delta_2) \mapsto g(\delta_1, \delta_2)$ exist and are continuous, hence the limit in (A.55) exists and is equal to

$$\begin{aligned}\lim_{\delta \to 0} \frac{F_a^{-1}(\tau - \delta U) - F_Y^{-1}(\tau)}{\delta} &= \left.\frac{\partial g(\delta_1, 0)}{\partial \delta_1}\right|_{\delta_1 = 0} + \left.\frac{\partial g(0, \delta_2)}{\partial \delta_2}\right|_{\delta_2 = 0} \\ &= -\frac{F_{Y|D=1}(F_Y^{-1}(\tau)) - F_{Y|D=0}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))} - \frac{U}{f_Y(F_Y^{-1}(\tau))}.\end{aligned}$$

For the upper bound, we have the analogous result

$$\lim_{\delta \to 0} \frac{F_a^{-1}(\tau - \delta L) - F_Y^{-1}(\tau)}{\delta} = -\frac{F_{Y|D=1}(F_Y^{-1}(\tau)) - F_{Y|D=0}(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))} - \frac{L}{f_Y(F_Y^{-1}(\tau))}.$$

□

Proof of Theorem 3.4. We introduce some new notation related to Assumption 3.5. Let $\mathbb{D}_\delta \subset \ell^\infty(\mathscr{Y})$ denote the set of all restrictions of distribution functions on $\mathbb{R}$ to $[F_Y^{-1}(\delta) - \varepsilon, F_Y^{-1}(1 - \delta) + \varepsilon]$. Additionally, $\mathbb{C}_\delta$ is the set of continuous functions on $[F_Y^{-1}(\delta) - \varepsilon, F_Y^{-1}(1 - \delta) + \varepsilon]$. Also, $\mathbb{UC}(\mathscr{Y})$ is the set of uniformly continuous functions defined on $\mathscr{Y}$.

The estimator of the apparent counterfactual distribution F_a is given by

$$\hat{F}_a(y) = (1 - \hat{p} - \delta)\hat{F}_{Y|D=0, D_\delta = 0}(y) + (\hat{p} + \delta)\hat{F}_{Y|D=1, D_\delta = 1}(y)$$

The apparent counterfactual can be written as the map $\mathbb{D}(\mathscr{Y})^2 \times (0,1) \mapsto \mathbb{D}(\mathscr{Y})$ given by

$$\begin{aligned}\psi(F_{Y|D=0,D_\delta=0}, F_{Y|D=1,D_\delta=1}, p) &= (1-p-\delta)F_{Y|D=0,D_\delta=0} + (p+\delta)F_{Y|D=1,D_\delta=1} \\ &= (1-\delta)F_{Y|D=0,D_\delta=0} + \delta F_{Y|D=1,D_\delta=1} \\ &+ (F_{Y|D=1,D_\delta=1} - F_{Y|D=0,D_\delta=0})p.\end{aligned}$$

This map is linear, so the Hadamard derivative tangentially to $\ell^\infty(\mathscr{Y})^2 \times (0,1)$ at $(F_{Y|D=0,D_\delta=0}, F_{Y|D=1,D_\delta=1}, p)$ is the map

$$\psi'_{F_{Y|D=0,D_\delta=0}, F_{Y|D=1,D_\delta=1}, p}(h_1, h_2, h_3) = (1-\delta)h_1 + \delta h_2 + (F_{Y|D=1,D_\delta=1} - F_{Y|D=0,D_\delta=0})h_3.$$

By the functional Delta method (see Theorem 20.8 in van der Vaart (1998)) and Assumption 3.4, we have

$$\begin{aligned}\sqrt{n}(\hat{F}_a - F_a) &= \sqrt{n}(\psi(\hat{F}_{Y|D=0,D_\delta=0}, \hat{F}_{Y|D=1,D_\delta=1}, \hat{p}) - \psi(F_{Y|D=0,D_\delta=0}, F_{Y|D=1,D_\delta=1}, p)) \\ &\rightsquigarrow \mathbb{G}_a := (1-\delta)\mathbb{G}_{0,0} + \delta\mathbb{G}_{1,1} + (F_{Y|D=1,D_\delta=1} - F_{Y|D=0,D_\delta=0})\mathbb{Z}_p.\end{aligned}$$

and convergence takes place in $\ell^\infty(\mathscr{Y})$. The random element $\mathbb{G}_a$ is Gaussian. Indeed, for any $y \in \mathscr{Y}$

$$\mathbb{G}_a(y) = (1-\delta)\mathbb{G}_{0,0}(y) + \delta\mathbb{G}_{1,1}(y) + (F_{Y|D=1,D_\delta=1}(y) - F_{Y|D=0,D_\delta=0}(y))\mathbb{Z}_p$$

is a linear combination of normal random variables.

Now we deal with

$$\sqrt{n}(\hat{\theta} - \theta) = -\frac{1}{\delta}\sqrt{n}\left(\hat{F}_a(\hat{F}_Y^{-1} + g) - F_a(F_Y^{-1} + g)\right)$$

This can be written as the composition of two maps. The first one is $\phi : \mathbb{D}(\mathscr{Y}) \times \mathbb{D}_\delta \mapsto$

$\mathbb{D}(\mathscr{Y}) \times \ell^\infty(\delta, 1-\delta)$ given by $\phi(H_1, H_2) \mapsto (H_1, H_2^{-1})$. The second one is $\psi : \mathbb{D}(\mathscr{Y}) \times \ell^\infty(\delta, 1-\delta) \mapsto \ell^\infty(\delta, 1-\delta)$ given by $\psi(H_1, H_2) \mapsto H_1 \circ (H_2 + g)$. Thus

$$\psi \circ \phi(F_a, F_Y) = F_a(F_Y^{-1} + g).$$

By Assumption 3.5 and Lemma 21.4(i) in van der Vaart (1998), ϕ has Hadamard derivative at (F_a, F_Y) tangentially to $\ell^\infty(\mathscr{Y}) \times \mathbb{C}_\delta$ given by the map

$$\phi'_{(F_a, F_Y)}(h_1, h_2) = \left(h_1, -\frac{h_2 \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}} \right).$$

The second map $\psi : \mathbb{D}(\mathscr{Y}) \times \ell^\infty(\delta, 1-\delta) \mapsto \ell^\infty(\delta, 1-\delta)$ is given by $\psi(H_1, H_2) \mapsto H_1 \circ (H_2 + g)$. It has Hadamard derivative tangentially to $\mathbb{UC}(\mathscr{Y}) \times \ell^\infty(\delta, 1-\delta)$ at any H_1 such that its derivative h_1 is bounded and uniformly continuous on $\mathscr{Y}$, and any H_2. To see, this we combine Lemmas 3.9.25 and 3.9.27 in van der Vaart and Wellner (1996). Let $\alpha_t \to \alpha$ and $\beta_t \to \beta$ in $\mathbb{D}(\mathscr{Y})$ and $\ell^\infty(\delta, 1-\delta)$ respectively, as $t \to 0$.

$$\begin{aligned}
&\frac{\psi(H_1 + t\alpha_t, H_2 + t\beta_t) - \psi(H_1, H_2)}{t} - \alpha \circ (H_2 + g) - h_1 \circ (H_2 + g) \cdot \beta \\
&= \frac{H_1 \circ (H_2 + g + t\beta_t) + t\alpha_t \circ (H_2 + g + t\beta_t) - H_1 \circ (H_2 + g)}{t} - \alpha \circ (H_2 + g) - h_1 \circ (H_2 + g) \cdot \beta \\
&= (\alpha_t - \alpha) \circ (H_2 + g + t\beta_t) + \alpha \circ (H_2 + g + t\beta_t) - \alpha \circ (H_2 + g) \\
&+ \frac{H_1 \circ (H_2 + g + t\beta_t) - H_1 \circ (H_2 + g)}{t} - h_1 \circ (H_2 + g) \cdot \beta
\end{aligned}$$

The first term, $(\alpha_t - \alpha) \circ (H_2 + g + t\beta_t)$, converges to 0 in $\mathbb{D}(\mathscr{Y})$ (that is, uniformly) because convergence of $\alpha_t \to \alpha$ is uniform. The second term, $\alpha \circ (H_2 + g + t\beta_t) - \alpha \circ (H_2 + g)$, converges to 0 in $\mathbb{D}(\mathscr{Y})$ because α is uniformly continuous on $\mathscr{Y}$. For the last term, fix a

$\tau \in (\delta, 1-\delta)$. By the mean-value theorem

$$\frac{H_1(H_2(\tau)+g+t\beta_t(\tau))-H_1(H_2(\tau)+g)}{t} - h_1(H_2(\tau)+g)\cdot\beta(\tau)$$
$$= h_1(\varepsilon_{\tau,t})\beta_t(\tau) - h_1(H_2(\tau)+g)\cdot\beta(\tau)$$
$$= h_1(\varepsilon_{\tau,t})(\beta_t(\tau)-\beta(\tau)) + (h_1(\varepsilon_{\tau,t}) - h_1(H_2(\tau)+g))\cdot\beta(\tau)$$

The first term, $h_1(\varepsilon_{\tau,t})(\beta_t(\tau)-\beta(\tau))$, converges uniformly to 0 because h_1 is bounded on $\mathscr{Y}$, and β_t converges uniformly to β. The second term converges to 0 uniformly because h_1 is uniformly continuous on $\mathscr{Y}$.

Hence, by Assumption 3.5, ψ has Hadamard derivative at (F_a, F_Y^{-1}) tangentially to $\mathbb{UC}(\mathscr{Y}) \times \ell^\infty(\delta, 1-\delta)$ given by the map

$$\psi'_{(F_a,F_Y^{-1})}(h_1,h_2) = h_1 \circ (F_Y^{-1}+g) + f_a \circ (F_Y^{-1}+g)\cdot h_2.$$

We use the chain rule (see Theorem 20.9 in van der Vaart (1998)) to conclude that $\psi \circ \phi$ has Hadamard derivative at (F_a, F_Y) tangentially to $\mathbb{UC}(\mathscr{Y}) \times \mathbb{C}_\delta$ given by the map

$$\begin{aligned}(\psi\circ\phi)'_{(F_a,F_Y)}(h_1,h_2) &= \psi'_{\phi(F_a,F_Y)} \circ \phi'_{(F_a,F_Y)}(h_1,h_2)\\ &= \psi'_{(F_a,F_Y^{-1})} \circ (h_1, -h_2(F_Y^{-1})/f_Y(F_Y^{-1}))\\ &= h_1 \circ (F_Y^{-1}+g) - f_a \circ (F_Y^{-1}+g)\frac{h_2 \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}}.\end{aligned}$$

By the functional Delta method (see Theorem 20.8 in van der Vaart (1998)) and the continuous mapping theorem (because of the $-1/\delta$ factor), we have that

$$\begin{aligned}\sqrt{n}(\hat{\theta}-\theta) &= -\frac{1}{\delta}\sqrt{n}\left(\hat{F}_a\circ(\hat{F}_Y^{-1}+g)-F_a\circ(F_Y^{-1}+g)\right)\\ &\rightsquigarrow -\frac{1}{\delta}(\psi\circ\phi)'_{(F_a,F_Y)}(\mathbb{G}_a,\mathbb{G}_Y))\\ &:= \mathbb{G}_\theta = -\frac{1}{\delta}\mathbb{G}_a\circ(F_Y^{-1}+g)+\frac{1}{\delta}f_a\circ(F_Y^{-1}+g)\frac{\mathbb{G}_Y\circ F_Y^{-1}}{f_Y\circ F_Y^{-1}}.\end{aligned}$$

To see that $\mathbb{G}_\theta$ is indeed Gaussian, we evaluate it at $\tau\in(\delta,1-\delta)$ to get

$$\mathbb{G}_\theta(\tau) = -\frac{1}{\delta}\mathbb{G}_a(F_Y^{-1}(\tau)+g)+\frac{1}{\delta}f_a(F_Y^{-1}(\tau)+g)\frac{\mathbb{G}_Y(F_Y^{-1}(\tau))}{f_Y(F_Y^{-1}(\tau))},$$

which is a linear combination of two normal random variables: $\mathbb{G}_a(F_Y^{-1}(\tau)+g)$ and $\mathbb{G}_Y(F_Y^{-1}(\tau))$. □

Proof of Theorem 3.5. The map given in (3.22) is

$$\phi(H) = \begin{pmatrix}\min\{\max\{0,H(\tau_1)\},1\}\\ \max\{\min\{0,H(\tau_2)\},-1\}\end{pmatrix}.$$

is the composition of an evaluation map $\theta\in\ell^\infty(\delta,1-\delta)\mapsto(\theta(\tau_1),\theta(\tau_2))$ and of the max/min composition. The evaluation map is linear, hence fully Hadamard differentiable. The composition of max/min is Hadamard directional differentiable by the chain rule for Hadamard directional differentiable maps (see Proposition 3.6 in Shapiro (1990); Lemma C2 of Masten and Poirier (2020)). Hence, another application of the chain rule yields that $\phi(H)$ is Hadamard directional differentiable at any $H\in\ell^\infty(\delta,1-\delta)$ tangentially to $\ell^\infty(\delta,1-\delta)$. By direct computation, the derivative, for any $h \in \ell^\infty(\delta, 1-\delta)$, is given by the map

$$\phi_H'(h) = \begin{pmatrix} h(\tau_1)\mathbb{1}_{\{0<H(\tau_1)<1\}} + \max(0, h(\tau_1))\mathbb{1}_{\{H(\tau_1)=0\}} + \min(0, h(\tau_1))\mathbb{1}_{\{H(\tau_1)=1\}} \\ h(\tau_2)\mathbb{1}_{\{-1<H(\tau_2)<0\}} + \min(0, h(\tau_2))\mathbb{1}_{\{H(\tau_2)=0\}} + \max(0, h(\tau_2))\mathbb{1}_{\{H(\tau_2)=-1\}} \end{pmatrix}. \tag{A.56}$$

Combining (A.56) with Theorem 2.1 in Fang and Santos (2019) and the result of Theorem 3.4, we arrive at

$$\sqrt{n}\begin{pmatrix} \hat{U}_{\tau_1} - U_{\tau_1} \\ \hat{L}_{\tau_2} - L_{\tau_2} \end{pmatrix} = \sqrt{n}(\phi(\hat{\theta}) - \phi(\theta)) \rightsquigarrow \phi_\theta'(\mathbb{G}_\theta),$$

where

$$\phi_\theta'(\mathbb{G}_\theta) = \begin{pmatrix} \mathbb{G}_\theta(\tau_1)\mathbb{1}_{\{0<\theta(\tau_1)<1\}} + \vee(0, \mathbb{G}_\theta(\tau_1))\mathbb{1}_{\{\theta(\tau_1)=0\}} + \wedge(0, \mathbb{G}_\theta(\tau_1))\mathbb{1}_{\{\theta(\tau_1)=1\}} \\ \mathbb{G}_\theta(\tau_2)\mathbb{1}_{\{-1<\theta(\tau_2)<0\}} + \wedge(0, \mathbb{G}_\theta(\tau_2))\mathbb{1}_{\{\theta(\tau_2)=0\}} + \vee(0, \mathbb{G}_\theta(\tau_2))\mathbb{1}_{\{\theta(\tau_2)=-1\}} \end{pmatrix}.$$

where $\wedge$ is the min operator and $\vee$ is the max operator.

□

Proof of Theorem 3.6. Recall that by (3.22)

$$\phi(\theta) = \begin{pmatrix} U_{\tau_1} \\ L_{\tau_2} \end{pmatrix},$$

This map is not fully differentiable with respect to θ, only directional differentiable. Now, for fixed τ, consider the map

$$\psi(F_a, F_Y, \theta, \tau) = \begin{pmatrix} F_a^{-1}(\tau - \delta\phi(\theta)_1) - F_Y^{-1}(\tau) \\ F_a^{-1}(\tau - \delta\phi(\theta)_2) - F_Y^{-1}(\tau) \end{pmatrix}, \tag{A.57}$$

where $\phi(\theta)_1$ and $\phi(\theta)_2$ are the first and second coordinates of $\phi(\theta)$ respectively. We want to find the distribution of

$$\sqrt{n}\left(\psi(\hat{F}_a, \hat{F}_Y, \hat{\theta}, \tau) - \psi(F_a, F_Y, \theta, \tau)\right)$$

Recall the notation introduced before: $\mathbb{D}_\delta \subset \ell^\infty(\mathscr{Y})$ denotes the set of all restrictions of distribution functions on $\mathbb{R}$ to $[F_Y^{-1}(\delta) - \varepsilon, F_Y^{-1}(1-\delta) + \varepsilon]$ for some $\varepsilon > 0$. Additionally, $\mathbb{C}_\delta$ is set of continuous functions on $[F_Y^{-1}(\delta) - \varepsilon, F_Y^{-1}(1-\delta) + \varepsilon]$.

Consider the map from $\mathbb{D}_\delta^2 \times \ell^\infty(\delta, 1-\delta) \mapsto \ell^\infty(\delta, 1-\delta)^2 \times [0,1] \times [-1,0]$ given by

$$m(H_1, H_2, H_3) = (H_1^{-1}, H_2^{-1}, \phi(H_3)_1, \phi(H_3)_2), \tag{A.58}$$

for ϕ defined in (3.22). Now consider the map from $\ell^\infty(\delta, 1-\delta)^2 \times [0,1] \times [-1,0] \mapsto \ell^\infty(\delta, 1-\delta)^2$ given by

$$q(H_1, H_2, H_3, H_4) = \begin{pmatrix} H_1(\cdot - \delta H_3) - H_2(\cdot) \\ H_1(\cdot - \delta H_4) - H_2(\cdot) \end{pmatrix}. \tag{A.59}$$

We can see that ψ in (A.57) is the composition

$$\psi(F_a, F_Y, \theta, \cdot) = q \circ m(F_a, F_Y, \theta).$$

By Assumptions 3.5 and 3.6, Lemma 21.4(i) in van der Vaart (1998), Theorem 3.4 and the chain rule for Hadamard directional differentiable maps, the map m is Hadamard directional differentiable (see Proposition 3.6 in Shapiro (1990); Lemma C2 of Masten and Poirier (2020)) at $(F_a, F_Y, \theta(\tau_1), \theta(\tau_2))$ tangentially to $\mathbb{C}_\delta^2 \times \ell^\infty(\delta, 1-\delta)$, with derivative given by the map

$$m'_{(F_a, F_Y, \theta)}(h_1, h_2, h_3) = \left(-\frac{h_1 \circ F_a^{-1}}{f_a \circ F_a^{-1}}, -\frac{h_2 \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}}, \phi'_\theta(h_3)_1, \phi'_\theta(h_3)_2\right) \tag{A.60}$$

where the map $h \mapsto \phi'_H(h)$ is given in (A.56), and $\phi'_H(h)_1$ and $\phi'_H(h)_2$ are the first and second coordinates respectively.

The map $q(H_1,H_2,H_3,H_4)$ in (A.59) has Hadamard derivative at $(F_a^{-1},F_Y^{-1},U_{\tau_1},L_{\tau_2})$ tangentially to $\mathbb{UC}(\delta,1-\delta)\times\ell^\infty(\delta,1-\delta)\times[0,1]\times[-1,0]$ given by the map

$$q'_{(F_a^{-1},F_Y^{-1},U_{\tau_1},L_{\tau_2})}(h_1,h_2,h_3,h_4) = \begin{pmatrix} h_1(\cdot-\delta U_{\tau_1}) - \frac{\delta h_4}{f_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})} - h_2(\cdot) \\ h_1(\cdot-\delta L_{\tau_2}) - \frac{\delta h_3}{f_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})} - h_2(\cdot) \end{pmatrix}.$$

We use the chain rule to conclude that the map $q\circ m$ has Hadamard directional derivative at (F_a,F_Y,θ) tangentially to $\mathbb{C}_\delta^2\times\ell^\infty(\delta,1-\delta)$ given by the map

$$\begin{aligned}(q\circ m)'_{(F_a,F_Y,\theta)}(h_1,h_2,h_3) &= q'_{m(F_a,F_Y,\theta)}\circ m'_{(F_a,F_Y,\theta)}(h_1,h_2,h_3)\\ &= q'_{(F_a^{-1},F_Y^{-1},L_{\tau_2},U_{\tau_1})}\\ &\circ\left(-\frac{h_1\circ F_a^{-1}}{f_a\circ F_a^{-1}},-\frac{h_2\circ F_Y^{-1}}{f_Y\circ F_Y^{-1}},\phi'_\theta(h_3)_1,\phi'_\theta(h_3)_2\right)\\ &= \begin{pmatrix} -\frac{h_1\circ F_a^{-1}(\cdot-\delta U_{\tau_1})}{f_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})} - \frac{\delta\phi'_\theta(h_3)_2}{f_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})} - \frac{h_2\circ F_Y^{-1}}{f_Y\circ F_Y^{-1}(\cdot)} \\ -\frac{h_1\circ F_a^{-1}(\cdot-\delta L_{\tau_2})}{f_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})} - \frac{\delta\phi'_\theta(h_3)_1}{f_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})} - \frac{h_2\circ F_Y^{-1}}{f_Y\circ F_Y^{-1}(\cdot)}\end{pmatrix}.\end{aligned}$$

Using Assumption 3.4, Theorem 3.4 and Theorem 2.1 in Fang and Santos (2019), we conclude that

$$\begin{aligned}\sqrt{n}\left(\psi(\hat{F}_a,\hat{F}_Y,\hat{\theta},\cdot)-\psi(F_a,F_Y,\theta,\cdot)\right) &\rightsquigarrow (q\circ m)'_{(F_a,F_Y,\theta)}(\mathbb{G}_a,\mathbb{G}_Y,\mathbb{G}_\theta)\\ &= \begin{pmatrix} -\frac{\mathbb{G}_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})}{f_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})} - \frac{\delta\phi'_\theta(\mathbb{G}_\theta)_2}{f_a\circ F_a^{-1}(\cdot-\delta U_{\tau_1})} - \frac{\mathbb{G}_Y\circ F_Y^{-1}(\cdot)}{f_Y\circ F_Y^{-1}(\cdot)} \\ -\frac{\mathbb{G}_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})}{f_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})} - \frac{\delta\phi'_\theta(\mathbb{G}_\theta)_1}{f_a\circ F_a^{-1}(\cdot-\delta L_{\tau_2})} - \frac{\mathbb{G}_Y\circ F_Y^{-1}(\cdot)}{f_Y\circ F_Y^{-1}(\cdot)}\end{pmatrix},\end{aligned}$$

and convergence takes place in $\ell^\infty(\delta,1-\delta)\times\ell^\infty(\delta,1-\delta)$.

□

Proof of Theorem 3.7. For $d = 0$ or $d = 1$, we find the asymptotic distribution of

$$\sqrt{n}(\hat{F}_{Y|D=d} \circ \hat{F}_Y^{-1} - F_{Y|D=d} \circ F_Y^{-1})$$

. Consider first the map $\psi : \mathbb{D}(\mathscr{Y})^2 \to \mathbb{D}(\mathscr{Y}) \times \ell^\infty(0,1)$, given by $\psi(H_1, H_2) = (H_1, H_2^{-1})$. Here, $\mathbb{D}(\mathscr{Y})$ is the set of all restrictions of distribution functions on $\mathbb{R}$ to $\mathscr{Y} = [y_l, y_u]$, such that they give mass 1 to $(y_l, y_u]$. Also, $\mathbb{C}(\mathscr{Y})$ is the set of all (uniformly) continuous functions defined on $\mathscr{Y}$.

By Lemma 21.4.(ii) in van der Vaart (1998), and Assumption 3.8, ψ is Hadamard differentiable tangentially to $\ell^\infty(\mathscr{Y}) \times \mathbb{C}(\mathscr{Y})$ at $(F_{Y|D=d}, F_Y)$, with derivative given by the map

$$\psi'_{(F_{Y|D=d}, F_Y)}(h_1, h_2) = \left(h_1, -\frac{h_2 \circ F_Y^{-1}}{f_y \circ F_Y^{-1}} \right).$$

Now, consider the map $\phi : \mathbb{D}(\mathscr{Y}) \times \ell^\infty(0,1) \to \ell^\infty(0,1)$ given by $\phi(H_1, H_2) = H_1 \circ H_2^{-1}$. By Lemmas 3.9.25 and 3.9.27 in van der Vaart and Wellner (1996), and Assumption 3.8, ϕ has Hadamard derivative at $(F_{Y|D=d}, F_Y^{-1})$ tangentially to $\mathbb{UC}(\mathscr{Y}) \times \ell^\infty(0,1)$ given by the map

$$\phi'_{(F_{Y|D=d}, F_Y^{-1})}(h_1, h_2) = h_1 \circ F_Y^{-1} + f_{Y|D=d} \circ F_Y^{-1} \cdot h_2.$$

We use the chain rule (see Theorem 20.9 in van der Vaart (1998)) to conclude that $\phi \circ \psi$ has Hadamard derivative at $(F_{Y|D=d}, F_Y)$ tangentially to $\mathbb{UC}(\mathscr{Y}) \times \mathbb{C}(\mathscr{Y})$ given by the map

$$\begin{aligned}
(\phi \circ \psi)'_{(F_{Y|D=d}, F_Y)}(h_1, h_2) &= \phi'_{\phi(F_{Y|D=d}, F_Y)} \circ \psi'_{(F_{Y|D=d}, F_Y)}(h_1, h_2) \\
&= \phi'_{(F_{Y|D=d}, F_Y^{-1})} \circ (h_1, -h_2 \circ F_Y^{-1} / f_Y \circ F_Y^{-1}) \\
&= h_1 \circ F_Y^{-1} - f_{Y|D=d} \circ F_Y^{-1} \cdot \frac{h_2 \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}}.
\end{aligned}$$

By the functional Delta method (see Theorem 20.8 in van der Vaart (1998)) we have that

$$\sqrt{n}(\hat{F}_{Y|D=d} \circ \hat{F}_Y^{-1} - F_{Y|D=d} \circ F_Y^{-1}) \rightsquigarrow (\phi \circ \psi)'_{(F_{Y|D=d}, F_Y)}(\mathbb{G}_d, \mathbb{G}_Y))$$

$$\mathbb{G}_{d,Y} := \mathbb{G}_d \circ F_Y^{-1} - f_{Y|D=d} \circ F_Y^{-1} \cdot \frac{\mathbb{G}_Y \circ F_Y^{-1}}{f_Y \circ F_Y^{-1}}.$$

By the continuous mapping theorem

$$\begin{aligned}\sqrt{n}(\hat{\theta} - \theta) &= \sqrt{n}\left(\hat{F}_{Y|D=0} \circ \hat{F}_Y^{-1} - \hat{F}_{Y|D=1} \circ \hat{F}_Y^{-1} - \left(F_{Y|D=0} \circ F_Y^{-1} - F_{Y|D=1} \circ F_Y^{-1}\right)\right) \\ &\rightsquigarrow \mathbb{G}_{0,Y} - \mathbb{G}_{1,Y}.\end{aligned}$$

$\square$

Bibliography

Beare, Brendan K. and Xiaoxia Shi (2019), "An Improved Bootstrap Test of Density Ratio Ordering." *Econometrics and Statistics*, 10, 9–26.

Bjorklund, Anders and Robert Moffitt (1987), "The Estimation of Wage Gains and Welfare Gains in Self-Selection." *The Review of Economics and Statistics*, 69, 42–49.

Card, David (1996), "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica*, 64, 957–979.

Card, David (2001), "The Effect of Unions on Wage Inequality in the U.S. Labor Market." *Industrial and Labor Relations Review*, 54, 296–315.

Card, David, Thomas Lemieux, and W. Craig Riddell (2004), "Unions and Wage Inequality." *Journal of Labor Research*, 25, 519–559.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil (2010), "Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin." *Econometrica*, 78, 377–394.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil (2011), "Estimating Marginal Returns to Education." *American Economic Review*, 101, 2754–2781.

Carneiro, Pedro and Sokbae Lee (2009), "Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality." *Journal of Econometrics*, 149, 191–208.

Dümbgen, L. (1993), "On Nondifferentiable Functions and the Bootstrap." *Probability Theory and Related Fields*, 95, 125–140.

Fang, Zheng and Andres Santos (2019), "Inference on Directionally Differentiable Functions." *The Review of Economic Studies*, 86, 377–412.

Farber, Henry S., Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu (2020), "Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data." Working Paper.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2007), "Unconditional Quantile Regressions." *NBER Technical Working Paper 339*.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2009), "Unconditional Quantile Regressions." *Econometrica*, 77, 953–973.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2011), "Decomposition Methods in Economics." *Handbook of Labor Economics*, 4, 1–102.

Freeman, Richard B. (1980), "Unionism and the Dispersion of Wages." *Industrial and Labor Relations Review*, 34, 3–23.

Hahn, Jinyong and Geert Ridder (2013), "Asymptotic variance of semiparametric estimators with generated regressors." *Econometrica*, 81, 315–340.

Heckman, James J. (1997), "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *The Journal of Human Resources*, 32, 441–462.

Heckman, James J. and Edward Vytlacil (1999), "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences of the United States of America*, 96, 4730–4734.

Heckman, James J. and Edward Vytlacil (2001a), "Local Instrumental Variables." In *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya* (C. Hsiao, K. Morimune, and J. Powell, eds.), 1–46, Cambrigde University Press, Cambridge, UK.

Heckman, James J. and Edward Vytlacil (2001b), "Policy Relevant Treatment Effects." *American Economic Review*, 91, 107–111.

Heckman, James J. and Edward Vytlacil (2005), "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica*, 73, 669–738.

Hong, Han and Jessie Li (2018), "The Numerical Delta Method." *Journal of Econometrics*, 206, 379–394.

Hong, Han and Jessie Li (2020), "The Numerical Bootstrap." *The Annals of Statistics*, 48, 397–412.

Hsu, Yu-Chin, Tsung-Chih Lai, and Robert P. Lieli (2020), "Counterfactual Treatment Effects: Estimation and Inference." *Journal of Business and Economic Statistics*, Forthcoming.

Ibragimov, I. A. and R. Z. Hasminskii (1981), *Statistical Estimation: Asymptotic Theory*. Springer-Verlag.

Imbens, Guido and Donald Rubin (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge.

Imbens, Guido W. and Whitney K. Newey (2009), “Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity.” *Econometrica*, 77, 1481–1512.

Kaplan, David M. (2020), “Assessing Policy Effects with Unconditional Quantile Regression.” Working Paper.

Kasy, Maximilian (2016), “Partial Identification, Distributional Preferences, and The Welfare Ranking of Policies.” *The Review of Economics and Statistics*, 98, 111–131.

Kline, Patrick and Andres Santos (2013), “Sensitivity to Missing Data Assumptions: Theory and an Evaluation of the U.S. Wage Structure.” *Quantitative Economics*, 4, 231–267.

Lemieux, Thomas (2006), “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review*, 96, 461–498.

Manski, Charles F (1989), “Anatomy of the Selection Problem.” *The Journal of Human Resources*, 24, 343–360.

Manski, Charles F (1990), “Nonparametric Bounds on Treatment Effects.” *American Economic Review Papers and Proceedings*, 80, 319–323.

Martinez-Iriarte, Julian and Yixiao Sun (2020), “Identification and Estimation of Unconditional Policy Effects of an Endogenous Binary Treatment.” Working Paper.

Masten, Matthew A. and Alexandre Poirier (2018), “Identification of Treatment Effects Under Conditional Partial Independence.” *Econometrica*, 86, 317–351.

Masten, Matthew A. and Alexandre Poirier (2020), “Inference on Breakdown Frontiers.” *Quantitative Economics*, 11, 41–111.

Newey, Whitney K. (1994), “The asymptotic variance of semiparametric estimators.” *Econometrica*, 62, 1349–1382.

Pagan, Adrian and Aman Ullah (1999), *Nonparametric Econometrics*. Cambridge University Press.

Rothe, Christoph (2010), “Identification of Unconditional Partial Effects in Nonseparable Models.” *Economics Letters*, 109, 171–174.

Rothe, Christoph (2012), “Partial Distributional Policy Effects.” *Econometrica*, 80, 2269–2301.

Rudin, Walter (1976), *Principles of Mathematical Analysis*. New York: McGraw-Hill.

Serfling, Robert J. (1980), *Approximation Theorems of Mathematical Statistics*. New York: Wiley.

Shapiro, A. (1990), "On Concepts of Directional Differentiability." *Journal of Optimization Theory and Applications*, 66, 477–487.

van der Vaart, A. W. (1998), *Asymptotic Statistics*. Cambridge University Press.

van der Vaart, A. W. and Jon A. Wellner (1996), *Weak Convergence and Empirical Processes*. Springer.

Zhou, Xiang and Yu Xie (2019), "Marginal Treatment Effects from a Propensity Score Perspective." *Journal of Political Economy*, 127, 3070–3084.