ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 86, No. 1 (January, 2018), 317–351

IDENTIFICATION OF TREATMENT EFFECTS UNDER CONDITIONAL PARTIAL INDEPENDENCE

MATTHEW A. MASTEN
Department of Economics, Duke University

ALEXANDRE POIRIER
Department of Economics, University of Iowa

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website http://www.econometricsociety.org or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

Econometrica, Vol. 86, No. 1 (January, 2018), 317–351

IDENTIFICATION OF TREATMENT EFFECTS UNDER CONDITIONAL PARTIAL INDEPENDENCE

MATTHEW A. MASTEN
Department of Economics, Duke University

ALEXANDRE POIRIER
Department of Economics, University of Iowa

Conditional independence of treatment assignment from potential outcomes is a commonly used but nonrefutable assumption. We derive identified sets for various treatment effect parameters under nonparametric deviations from this conditional independence assumption. These deviations are defined via a conditional treatment assignment probability, which makes it straightforward to interpret. Our results can be used to assess the robustness of empirical conclusions obtained under the baseline conditional independence assumption.

KEYWORDS: Treatment effects, conditional independence, unconfoundedness, selection on observables, sensitivity analysis, nonparametric identification, partial identification.

1. INTRODUCTION

THE TREATMENT EFFECT MODEL UNDER CONDITIONAL INDEPENDENCE is widely used in empirical research. The conditional independence assumption states that, after conditioning on a set of observed covariates, treatment assignment is independent of potential outcomes. This assumption has many other names, including unconfoundedness, ignorability, exogenous selection, and selection on observables. It delivers point identification of many parameters of interest, including the average treatment effect, the average effect of treatment on the treated, and quantile treatment effects. Imbens and Rubin (2015) provided a recent overview of this literature.

Without additional data, like the availability of an instrument, the conditional independence assumption is not refutable: The data alone cannot tell us whether the assumption is true. Moreover, conditional independence is often considered a strong and controversial assumption. Consequently, empirical researchers may wonder: How credible are treatment effect estimates obtained under conditional independence?

In this paper, we address this concern by studying what can be learned about treatment effects under a nonparametric class of assumptions that are weaker than conditional independence. While there are many ways to weaken independence, we focus on just one, which we call *conditional c-dependence*.[1] This assumption states that the probability of being treated given observed covariates and an unobserved potential outcome is not too far from the probability of being treated given just the observed covariates. We use the sup-norm distance, where the scalar c denotes how much these two conditional probabilities

Matthew A. Masten: matt.masten@duke.edu
Alexandre Poirier: alexandre-poirier@uiowa.edu

This paper is based on portions of our previous working papers, Masten and Poirier (2016, 2017). We thank the editor, the referees, audiences at various seminars, as well as Federico Bugni, Ivan Canay, Joachim Freyberger, Joe Hotz, Guido Imbens, Shakeeb Khan, Chuck Manski, Jim Powell, Adam Rosen, Suyong Song, and Alex Torgovitsky, for helpful conversations and comments.

[1]See Masten and Poirier (2016) for an analysis and discussion of several other approaches. We refer to any of these approaches as *partial* independence assumptions.

© 2018 The Econometric Society https://doi.org/10.3982/ECTA14481

may differ. This class of assumptions nests both the conditional independence assumption and the opposite end of no constraints on treatment selection.

In our first main contribution, we derive sharp bounds on conditional c.d.f.s that are consistent with conditional c-dependence. This result can be used in many models, including the treatment effects model we study here.[2] In that model, as our second main contribution, we derive identified sets for many parameters of interest. These include the average treatment effect, the average effect of treatment on the treated, and quantile treatment effects. These identified sets have simple, analytical characterizations. Empirical researchers can use these identified sets to examine how sensitive their parameter estimates are to deviations from the baseline assumption of conditional independence. We illustrate this sensitivity analysis in a brief numerical example.

Related Literature

In the rest of this section, we review the related literature. As discussed in Section 22.4 of Imbens and Rubin (2015), a large literature starting with the seminal work of Rosenbaum and Rubin (1983) relaxes conditional independence by modeling the conditional probabilities of treatment assignment given both observable and unobservable variables parametrically. This literature also typically imposes a parametric model on outcomes. This includes Lin, Psaty, and Kronmal (1998), Imbens (2003), and Altonji, Elder, and Taber (2005, 2008). An important exception is Robins, Rotnitzky, and Scharfstein (2000), who relaxed parametric assumptions on outcomes. They continued to use parametric models for treatment assignment probabilities, however, when applying their results. Our work builds on this literature by developing fully nonparametric methods for sensitivity analysis. Our new methods can ensure that empirical findings of robustness do not rely on auxiliary parametric assumptions.

We are aware of only two previous analyses in this sensitivity analysis literature which develop fully nonparametric methods. The first is Ichino, Mealli, and Nannicini (2008), who avoided specifying a parametric model by assuming that all observed and unobserved variables are discretely distributed, so that their joint distribution is determined by a finite dimensional vector. Unlike our approach, theirs rules out continuous outcomes. It also involves many different sensitivity parameters, while our approach uses only one sensitivity parameter.

The second is Rosenbaum (1995, 2002a), who proposed a sensitivity analysis within the context of randomization inference for testing the sharp null hypotheses of no unit level treatment effects for all units in one's data set. Since this approach is based on finite sample randomization inference (cf., Chapter 5 of Imbens and Rubin (2015)), rather than population level identification analysis, this is quite different from the approaches discussed above and from what we do in the present paper. Like our results, however, his approach does not impose a parametric model on treatment assignment probabilities.

A large literature initiated by Manski has studied identification problems under various assumptions which typically do not point identify the parameters (e.g., Manski (2007)). In the context of missing data analysis, Manski (2016) suggested imposing a class of assumptions which includes conditional c-dependence. He did not, however, derive any identified sets under this assumption. Several papers study partial identification of treatment response under deviations from mean independence assumptions, rather than the

[2]See Masten and Poirier (2016) for several other applications of this result.

statistical independence assumption we start from. Manski and Pepper (2000, 2009) relaxed mean independence to a monotonicity constraint in the conditioning variable, while Hotz, Mullin, and Sanders (1997) supposed mean independence only holds for some portion of the population. These relaxations and conditional c-dependence are non-nested. Moreover, these papers focus on mean potential outcomes, while we also study quantiles and distribution functions. Finally, Manski's original no assumptions bounds for average treatment effects (Manski 1989, 1990) are obtained as a special case of our conditional c-dependence ATE bounds when c is sufficiently large.

2. MODEL, ASSUMPTIONS, AND INTERPRETATION

We study the standard potential outcomes model with a binary treatment. In this section we set up the notation and some maintained assumptions. We define our parameters of interest and state the key assumption which point identifies them: random assignment of treatment, conditional on covariates. We discuss how we relax this assumption. We derive identified sets under these relaxations in Section 3. We conclude this section by suggesting a few ways to interpret our deviations from conditional independence.

Basic Setup

Let Y be an observed scalar outcome variable and $X \in \{0, 1\}$ an observed binary treatment. Let Y_1 and Y_0 denote the unobserved potential outcomes. As usual, the observed outcome is related to the potential outcomes via the equation

$$Y = XY_1 + (1 - X)Y_0. \tag{1}$$

Let $W \in \text{supp}(W)$ denote a vector of observed covariates, which may be discrete, continuous, or mixed. Let $p_{x|w} = \mathbb{P}(X = x \mid W = w)$ denote the observed generalized propensity score (Imbens (2000)). We consider both continuous and binary potential outcomes. We begin with the continuous outcome case, where we maintain the following assumption on the joint distribution of (Y_1, Y_0, X, W).

ASSUMPTION A1: For each $x, x' \in \{0, 1\}$ and $w \in \text{supp}(W)$:

1. $Y_x \mid X = x', W = w$ has a strictly increasing and continuous distribution function on $\text{supp}(Y_x \mid X = x', W = w)$.
2. $\text{supp}(Y_x \mid X = x', W = w) = \text{supp}(Y_x \mid W = w) = [\underline{y}_x(w), \overline{y}_x(w)]$ where $-\infty \leq \underline{y}_x(w) < \overline{y}_x(w) \leq \infty$.
3. $p_{1|w} \in (0, 1)$ for all $w \in \text{supp}(W)$.

By equation (1),

$$F_{Y|X,W}(y \mid x, w) = \mathbb{P}(Y_x \leq y \mid X = x, W = w)$$

and hence Assumption A1.1 implies that the distribution function of $Y \mid X = x, W = w$ is also strictly increasing and continuous. By the law of iterated expectations, the marginal distributions of Y and Y_x have the same properties as well. We consider the binary outcome case on page 329.

Assumption A1.2 states that the conditional support of Y_x given $X = x', W = w$ does not depend on x', and that this support is a possibly infinite closed interval. The first equality is a "support independence" assumption, which is implied by the standard conditional independence assumption. Since $Y \mid X = x, W = w$ has the same distribution as $Y_x \mid X = x, W = w$, this implies that the support of $Y \mid X = x, W = w$ equals that of $Y_x \mid W = w$. Consequently, the endpoints $\underline{y}_x(w)$ and $\overline{y}_x(w)$ are point identified. Assumption A1.3 is a standard overlap assumption.

Define the *conditional rank* random variables $R_1 = F_{Y_1|W}(Y_1 \mid W)$ and $R_0 = F_{Y_0|W}(Y_0 \mid W)$. For any $w \in \operatorname{supp}(W)$, $R_1 \mid W = w$ and $R_0 \mid W = w$ are uniformly distributed on [0, 1], since $F_{Y_1|W}(\cdot \mid w)$ and $F_{Y_0|W}(\cdot \mid w)$ are strictly increasing. Moreover, by construction, both R_1 and R_0 are independent of W. The value of unit i's conditional rank random variable R_x tells us where unit i lies in the conditional distribution of $Y_x \mid W = w$. We occasionally use these variables throughout the paper.

Identifying Assumptions

It is well known that the conditional distributions of potential outcomes $Y_1 \mid W$ and $Y_0 \mid W$ and therefore the marginal distributions of Y_1 and Y_0 are point identified under the following assumption:

CONDITIONAL INDEPENDENCE: $X \perp\!\!\!\perp Y_1 \mid W$ *and* $X \perp\!\!\!\perp Y_0 \mid W$.

These marginal distributions are immediately point identified from

$$F_{Y_x|W}(y \mid w) = F_{Y|X,W}(y \mid x, w)$$

and

$$F_{Y_x}(y) = \int_{\operatorname{supp}(W)} F_{Y|X,W}(y \mid x, w)\, dF_W(w).$$

Consequently, any functional of $F_{Y_1|W}$ and $F_{Y_0|W}$ is also point identified under the conditional independence assumption. Leading examples include the average treatment effect,

$$\text{ATE} = \mathbb{E}(Y_1 - Y_0),$$

and the τth quantile treatment effect,

$$\text{QTE}(\tau) = Q_{Y_1}(\tau) - Q_{Y_0}(\tau),$$

where $\tau \in (0, 1)$. The goal of our identification analysis is to study what can be said about such functionals when conditional independence partially fails. To do this we define the following class of assumptions, which we call *conditional c-dependence*.

DEFINITION 1: Let $x \in \{0, 1\}$. Let $w \in \operatorname{supp}(W)$. Let c be a scalar between 0 and 1. Say X is *conditionally c-dependent* with Y_x given $W = w$ if

$$\sup_{y_x \in \operatorname{supp}(Y_x|W=w)} \left|\mathbb{P}(X = 1 \mid Y_x = y_x, W = w) - \mathbb{P}(X = 1 \mid W = w)\right| \leq c. \tag{2}$$

If (2) holds for all $w \in \operatorname{supp}(W)$, we say X is conditionally c-dependent with Y_x given W.

Under the conditional independence assumption $X \perp\!\!\!\perp Y_x \mid W$,

$$\mathbb{P}(X = 1 \mid Y_x = y_x, W = w) = \mathbb{P}(X = 1 \mid W = w)$$

for all $y_x \in \text{supp}(Y_x \mid W = w)$ and all $w \in \text{supp}(W)$. Conditional c-dependence allows for deviations from this assumption by allowing the conditional probability $\mathbb{P}(X = 1 \mid Y_x = y_x, W = w)$ to be different from the propensity score $p_{1|w}$, but not too different. This class of assumptions nests conditional independence as the special case where $c = 0$. Moreover, when $c \geq \max\{p_{1|w}, p_{0|w}\}$, from (2) we see that conditional c-dependence imposes no constraints on $\mathbb{P}(X = 1 \mid Y_x = y_x, W = w)$. Values of c strictly between zero and $\max\{p_{1|w}, p_{0|w}\}$ lead to intermediate cases. These intermediate cases can be thought of as a kind of limited selection on unobservables, since the value of one's unobserved potential outcome Y_x is allowed to affect the probability of receiving treatment.

Beginning with Rosenbaum and Rubin (1983), many papers use parametric models for unobserved conditional probabilities similar to $\mathbb{P}(X = 1 \mid Y_x = y_x, W = w)$ to model deviations from conditional independence. For example, see Robins, Rotnitzky, and Scharfstein (2000) and Imbens (2003). In contrast, conditional c-dependence is a nonparametric class of assumptions. Our results therefore ensure that empirical findings of robustness do not depend on any auxiliary parametric assumptions.

By invertibility of $F_{Y_x|W}(\cdot \mid w)$ for each $x \in \{0, 1\}$ and $w \in \text{supp}(W)$ (Assumption A1.1), equation (2) is equivalent to

$$\sup_{r_x \in [0,1]} \left| \mathbb{P}(X = 1 \mid R_x = r_x, W = w) - \mathbb{P}(X = 1 \mid W = w) \right| \leq c. \tag{2'}$$

Using this result, we obtain the following characterization of conditional c-dependence.

PROPOSITION 1: *Suppose Assumption* A1.1 *holds. Then X is conditionally c-dependent with the potential outcome Y_x given $W = w$ if and only if*

$$\sup_{x' \in \{0,1\}} \sup_{r \in [0,1]} \left| f_{X,R_x|W}(x', r \mid w) - p_{x'|w} f_{R_x|W}(r \mid w) \right| \leq c. \tag{3}$$

Proposition 1 shows that conditional c-dependence is equivalent to a constraint on the sup-norm distance between the joint p.d.f. of $(X, R_x) \mid W$ and the product of the marginal distributions of $X \mid W$ and $R_x \mid W$. Although we do not pursue this here, this alternative characterization also suggests how to extend this concept to continuous treatments. Finally, we note that another equivalent characterization of conditional c-dependence obtains by replacing $X = 1$ with $X = 0$ in equation (2).

Throughout the rest of the paper we impose conditional c-dependence between X and the potential outcomes given covariates:

ASSUMPTION A2: X is conditionally c-dependent with Y_1 given W and with Y_0 given W.

Interpreting Conditional c-Dependence

Interpreting the deviations from one's baseline assumption is an important part of any sensitivity analysis. In this subsection, we give several suggestions for how to interpret our sensitivity parameter c in practice.

Our first suggestion, going back to the earliest sensitivity analysis of Cornfield et al. (1959), and used more recently in Imbens (2003), Altonji, Elder, and Taber (2005, 2008), and Oster (2018), is to use the amount of selection on observables to calibrate our beliefs about the amount of selection on unobservables. To formalize this idea in our context, recall that conditional c-dependence is defined using a distance between two conditional treatment probabilities: the usual propensity score $\mathbb{P}(X = 1 \mid W = w)$ and that same conditional probability, except also conditional on an unobserved potential outcome Y_x. Hence the question is: How much does adding this extra conditioning variable affect the conditional treatment probability?

In the data, Y_x is unobserved, so we cannot answer this question directly. But we can examine the impact of adding additional observed covariates on conditional treatment probabilities (assuming $K \equiv \dim(W) \geq 1$). Specifically, suppose we partition our vector of covariates W into (W_{-k}, W_k), where W_k is the kth component of W and W_{-k} is a vector of the remaining $K-1$ components. Define

$$\overline{c}_k = \sup_{w_{-k}} \sup_{w_k} \left| \mathbb{P}\big(X = 1 \mid W = (w_{-k}, w_k)\big) - \mathbb{P}(X = 1 \mid W_{-k} = w_{-k}) \right|,$$

where we take suprema over $w_k \in \text{supp}(W_k \mid W_{-k} = w_{-k})$ and $w_{-k} \in \text{supp}(W_{-k})$. $\overline{c}_k$ tells us the largest amount that the observed conditional treatment probabilities with and without the variable W_k can differ.[3] Less formally, it is a measure of the marginal impact of including the kth variable on treatment assignment, given that we have already included the vector W_{-k}. Similarly to Cornfield et al. (1959) and the subsequent literature, the idea is that if adding an extra observed variable creates variation $\overline{c}_k$, then it might be reasonable to expect that adding the unobserved variable Y_x to our conditioning set may also create variation $\overline{c}_k$. In practice, one can compute and examine $\overline{c}_k$ for each k.

Our next suggestion is a variation which incorporates information on the distribution of W. Define

$$p_{1|W}(w_{-k}, w_k) = \mathbb{P}\big(X = 1 \mid W = (w_{-k}, w_k)\big)$$

and

$$p_{1|W_{-k}}(w_{-k}) = \mathbb{P}(X = 1 \mid W_{-k} = w_{-k}).$$

Rather than examining the largest point in the support of the random variable

$$\left| p_{1|W}(W_{-k}, W_k) - p_{1|W_{-k}}(W_{-k}) \right|,$$

we could also consider quantiles of this distribution, such as the 50th, 75th, or 90th percentiles. One could also plot the distribution of this random variable for each k.

These suggestions are a kind of nonparametric version of the implicit partial R^2's used by Imbens (2003) in his parametric model, or of the logit coefficients used by Rosenbaum and Rubin (1983). The overall idea is the same: We are trying to measure the partial effect of adding an extra conditioning covariate on the conditional probability of treatment.

Our final suggestion reiterates a point made by Rosenbaum (2002b, Section 7): Precise quantitative interpretations of sensitivity parameters like c are not always necessary. We can perform qualitative comparisons of robustness across different studies and data sets by comparing the corresponding bound functions, as we do in our numerical illustration on page 330. Imbens (2003) made a similar point, stating that "not ... all evaluations are equally sensitive to departures from the exogeneity assumption" (page 126). Such rankings of studies in terms of their robustness may help one aggregate findings across different studies. We leave a formal study of this kind of robustness-adjusted meta-analysis to future work.

[3]If $K = 1$, then one can compare $\mathbb{P}(X = 1 \mid W = w)$ with $\mathbb{P}(X = 1)$.

3. IDENTIFICATION UNDER CONDITIONAL c-DEPENDENCE

In this section, we study identification of treatment effects under conditional c-dependence. To do so, we start by deriving bounds on c.d.f.s under generic c-dependence. We then apply these results to obtain sharp bounds on various treatment effect functionals.

3.1. *Partial Identification of c.d.f.s*

In this subsection, we consider the relationship between a generic scalar random variable U and a binary variable $X \in \{0, 1\}$. We derive sharp bounds on the conditional c.d.f. of U given X when (a) the marginal distributions of U and X are known and (b) X is c-dependent with U, meaning that

$$\sup_{u \in \operatorname{supp}(U)} \left|\mathbb{P}(X = 1 \mid U = u) - \mathbb{P}(X = 1)\right| \leq c. \tag{4}$$

In the next subsection we will condition on W and apply this general result with $U = R_x$, the conditional rank variable to obtain sharp bounds on various treatment effect parameters.

Let $F_{U|X}(u \mid x) = \mathbb{P}(U \leq u \mid X = x)$ denote the unknown conditional c.d.f. of U given $X = x$. Let $F_U(u) = \mathbb{P}(U \leq u)$ denote the known marginal c.d.f. of U. Let $p_x = \mathbb{P}(X = x)$ denote the known marginal probability mass function of X.

Define

$$\overline{F}^c_{U|X}(u \mid x) = \min\left\{F_U(u) + \frac{c}{p_x} \min\{F_U(u), 1 - F_U(u)\}, \frac{F_U(u)}{p_x}, 1\right\}$$

and

$$\underline{F}^c_{U|X}(u \mid x) = \max\left\{F_U(u) - \frac{c}{p_x} \min\{F_U(u), 1 - F_U(u)\}, \frac{F_U(u) - 1}{p_x} + 1, 0\right\}.$$

THEOREM 1: *Suppose the following hold*:
1. *The marginal distributions of U and X are known*.
2. *U is continuously distributed*.
3. *Equation* (4) *holds*.
4. $p_1 \in (0, 1)$.

Let $\mathcal{F}_{\operatorname{supp}(U)}$ denote the set of all c.d.f.s on $\operatorname{supp}(U)$. *Then*, *for each $x \in \{0, 1\}$*, $F_{U|X}(\cdot \mid x) \in \mathcal{F}^c_{U|x}$, *where*

$$\mathcal{F}^c_{U|x} = \left\{F \in \mathcal{F}_{\operatorname{supp}(U)} : \underline{F}^c_{U|X}(u \mid x) \leq F(u) \leq \overline{F}^c_{U|X}(u \mid x) \text{ for all } u \in \operatorname{supp}(U)\right\}.$$

Furthermore, *for each* $\varepsilon \in [0, 1]$, *there exists a joint distribution of* (U, X) *consistent with assumptions* 1–4 *above and such that*

$$\begin{aligned} &\big(\mathbb{P}(U \leq u \mid X = 1), \mathbb{P}(U \leq u \mid X = 0)\big) \\ &\quad = \big(\varepsilon \underline{F}^c_{U|X}(u \mid 1) + (1 - \varepsilon)\overline{F}^c_{U|X}(u \mid 1), (1 - \varepsilon)\underline{F}^c_{U|X}(u \mid 0) + \varepsilon \overline{F}^c_{U|X}(u \mid 0)\big) \end{aligned} \tag{5}$$

for all $u \in \operatorname{supp}(U)$. *Consequently*, *for any $x \in \{0, 1\}$ and $u \in \operatorname{supp}(U)$*, *the pointwise bounds*

$$F_{U|X}(u \mid x) \in \left[\underline{F}^c_{U|X}(u \mid x), \overline{F}^c_{U|X}(u \mid x)\right]$$

are sharp.

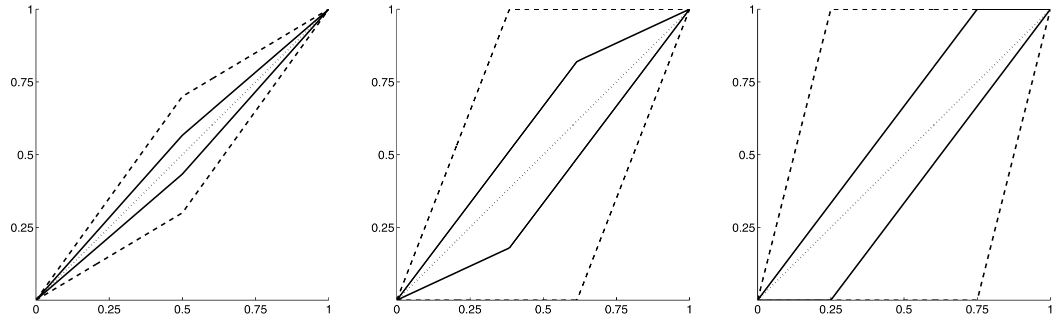

FIGURE 1.—Example upper and lower bounds on $F_{U|X}(u \mid x)$ for $x = 1$ (solid) and $x = 0$ (dashed), when $p_1 = 0.75$. Left: $c = 0.1 < \min\{p_1, p_0\}$. Middle: $\min\{p_1, p_0\} < c = 0.4 < \max\{p_1, p_0\}$. Right: $c = 0.9 > \max\{p_1, p_0\}$. The diagonal, representing the $c = 0$ case of full independence, is plotted as a dotted line.

The proof of Theorem 1, along with all other proofs, is given in the Appendix. In this proof, we note that there are two constraints on the conditional distribution of $U \mid X$. The first is the c-dependence constraint. The second is the fact that the marginal distributions of U and X are known, and hence the conditional c.d.f.s must satisfy a law of total probability constraint. This result is therefore a variation on the decomposition of mixtures problem. See Cross and Manski (2002), Manski (2007, Chapter 5), and Molinari and Peski (2006) for further discussion.

Theorem 1 has three conclusions. First, we show that the functions $\overline{F}^c_{U|X}(\cdot \mid x)$ and $\underline{F}^c_{U|X}(\cdot \mid x)$ bound the unknown conditional c.d.f. $F_{U|X}(\cdot \mid x)$ uniformly in their arguments. Second, we show that these bounds are functionally sharp in the sense that the joint identified set for the two conditional c.d.f.s $(F_{U|X}(\cdot \mid 1), F_{U|X}(\cdot \mid 0))$ contains linear combinations of the bound functions $\overline{F}^c_{U|X}(\cdot \mid x)$ and $\underline{F}^c_{U|X}(\cdot \mid x)$. Finally, we remark that this functional sharpness implies pointwise sharpness.

Importantly, the bound functions $\overline{F}^c_{U|X}(\cdot \mid x)$ and $\underline{F}^c_{U|X}(\cdot \mid x)$ are piecewise linear functions with simple analytical expressions. These bounds are proper c.d.f.s and can be attained, as stated above. As c approaches zero, these bounds for $F_{U|X}(u \mid x)$ collapse to the conditional c.d.f. $F_U(u)$. When c exceeds $\max\{p_0, p_1\}$, the c-dependence constraint is not binding. Consequently, the c.d.f. bounds simplify to

$$\overline{F}^c_{U|X}(u \mid x) = \min\left\{\frac{F_U(u)}{p_x}, 1\right\} \quad \text{and} \quad \underline{F}^c_{U|X}(u \mid x) = \max\left\{\frac{F_U(u) - 1}{p_x} + 1, 0\right\}.$$

These bounds can be interpreted as the *no assumptions bounds* since the only constraint imposed on the c.d.f.s is that they satisfy the law of total probability.

Figure 1 shows several examples of the bound functions $\overline{F}^c_{U|X}(\cdot \mid x)$ and $\underline{F}^c_{U|X}(\cdot \mid x)$. In this example we let $U \sim \text{Unif}[0, 1]$ and $p_1 = 0.75$. We let $c = 0.1$, 0.4, and 0.9, which represent three qualitative regions for c: (a) $c < \min\{p_1, p_0\}$, where both bounds are strictly between 0 and 1 on the interior of the support, (b) $\min\{p_1, p_0\} \leq c < \max\{p_1, p_0\}$, where one bound is strictly between 0 and 1 on the interior of the support, but the other is not, and (c) $c \geq \max\{p_1, p_0\}$, where we simply obtain the no assumptions bounds.

3.2. *Partial Identification of Treatment Effects*

Next we study identification of various treatment effects under conditional c-dependence. Throughout most of this section we focus on continuous outcomes. We study binary outcomes on page 329. We end with a numerical illustration of the identified set for ATE as a function of c, and discuss how this set depends on features of the observed distribution of the data.

Conditional c.d.f.s

Under conditional independence, $c = 0$, the marginal conditional distribution functions $F_{Y_0|W}$ and $F_{Y_1|W}$ are point identified. For $c > 0$, these functions are partially identified. In this case, we derive sharp bounds on these c.d.f.s in Proposition 2 below.

Define

$$\overline{F}^c_{Y_x|W}(y \mid w) = \min\left\{ \frac{p_{x|w}F_{Y|X,W}(y \mid x, w)}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c) + \mathbb{1}(p_{x|w} \le c), \right. \\ \left. \frac{p_{x|w}F_{Y|X,W}(y \mid x, w) + c}{p_{x|w} + c}, p_{x|w}F_{Y|X,W}(y \mid x, w) + (1 - p_{x|w}) \right\} \tag{6}$$

and

$$\underline{F}^c_{Y_x|W}(y \mid w) = \max\left\{ \frac{p_{x|w}F_{Y|X,W}(y \mid x, w)}{p_{x|w} + c}, \right. \\ \left. \frac{p_{x|w}F_{Y|X,W}(y \mid x, w) - c}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c), p_{x|w}F_{Y|X,W}(y \mid x, w) \right\} \tag{7}$$

for all $y \in [\underline{y}_x(w), \overline{y}_x(w))$. For $y < \underline{y}_x(w)$, define these c.d.f. bounds to be 0. For $y \ge \overline{y}_x(w)$, define these c.d.f. bounds to be 1.

The following proposition shows that the functions (6) and (7) are sharp bounds on the c.d.f. of $Y_x \mid W = w$ under conditional c-dependence.

PROPOSITION 2—Bounds on Conditional c.d.f.s: *Let* $w \in \operatorname{supp}(W)$. *Suppose the joint distribution of* (Y, X, W) *is known. Let Assumptions A1 and A2 hold. Let* $\mathcal{F}_{\mathbb{R}}$ *denote the set of all c.d.f.s on* $\mathbb{R}$.

1. *Then* $F_{Y_x|W}(\cdot \mid w) \in \mathcal{F}^c_{Y_x|w}$ *where*

$$\mathcal{F}^c_{Y_x|w} = \left\{ F \in \mathcal{F}_{\mathbb{R}} : \underline{F}^c_{Y_x|W}(y \mid w) \le F(y) \le \overline{F}^c_{Y_x|W}(y \mid w) \text{ for all } y \in \mathbb{R} \right\}.$$

2. *Furthermore, for each* $\varepsilon \in [0, 1]$ *and* $0 < \eta < \min\{p_{x|w}, 1 - p_{x|w}\}$, *there exist c.d.f.s* $F^c_{Y_x|W}(\cdot \mid w; \varepsilon, \eta) \in \mathcal{F}^c_{Y_x|w}$ *for* $x \in \{0, 1\}$ *such that*

(a) *There exists a joint distribution of* $(Y_1, Y_0, X) \mid W = w$ *consistent with our maintained assumptions and such that*

$$\mathbb{P}(Y_1 \le y \mid W = w) = F^c_{Y_1|W}(y \mid w; \varepsilon, \eta) \quad \text{and}$$
$$\mathbb{P}(Y_0 \le y \mid W = w) = F^c_{Y_0|W}(y \mid w; 1 - \varepsilon, \eta)$$

for all $y \in \operatorname{supp}(Y_x \mid W = w)$.

(b) *For each* $x \in \{0, 1\}$, *as* $\eta \searrow 0$, $F^c_{Y_x|W}(\cdot \mid w; 1, \eta)$ *and* $F^c_{Y_x|W}(\cdot \mid w; 0, \eta)$ *converge pointwise monotonically to* $\overline{F}^c_{Y_x|W}(\cdot \mid w)$ *and* $\underline{F}^c_{Y_x|W}(\cdot \mid w)$, *respectively.*

(c) *For each* $y \in \mathbb{R}$ *and* $w \in \text{supp}(W)$, *the function* $F^c_{Y_x|W}(y \mid w; \cdot, \eta)$ *is continuous on* $[0, 1]$.

3. *Consequently, for any* $y \in \mathbb{R}$, *the pointwise bounds*

$$F_{Y_x|W}(y \mid w) \in \big[\underline{F}^c_{Y_x|W}(y \mid w), \overline{F}^c_{Y_x|W}(y \mid w)\big] \tag{8}$$

have a sharp interior.

The proof of this result relies importantly on our general result, Theorem 1. Similarly to that result, Proposition 2 has three conclusions. First, we show that the functions (6) and (7) bound $F_{Y_x|W}(\cdot \mid w)$ uniformly in their arguments. Second, we show that these bounds are functionally sharp. As in Theorem 1, sharpness is subtle because $\mathcal{F}^c_{Y_x|w}$ is not the identified set for $F_{Y_x|W}(\cdot \mid w)$—it contains some c.d.f.s which cannot be attained. For example, it contains c.d.f.s with jump discontinuities on their support, which violates Assumption A1.1. We could impose extra constraints to obtain the sharp set of c.d.f.s, but this is not required for our analysis.

That said, the bound functions (6) and (7) used to define $\mathcal{F}^c_{Y_x|W}$ *are* sharp for the function $F_{Y_x|W}(\cdot \mid w)$ in the sense that there are c.d.f.s $F_{Y_x|W}(\cdot \mid w; \varepsilon, \eta)$ which (a) are attainable, (b) can be made arbitrarily close to the bound functions, and (c) continuously vary between the lower and upper bounds. The bound functions themselves are not always continuous and so violate Assumption A1.1. This explains the presence of the η variable. It also explains why the endpoints of the bounds in our results below may not be attainable. If c is small enough, the bound functions can be attainable, but we do not enumerate these cases for simplicity.

We use attainability of the functions $F_{Y_x|W}(\cdot \mid w; \varepsilon, \eta)$ to prove sharpness of identified sets for various functionals of $F_{Y_1|W}$ and $F_{Y_0|W}$. For example, the third conclusion to Proposition 2 states a pointwise-in-y sharpness result for the evaluation functional. In general, we obtain bounds for functionals by evaluating the functional at the bounds (6) and (7). Sharpness of these bounds then follows by applying Proposition 2.

Conditional QTEs, CATE, and ATE

Next we derive identified sets for functionals of the marginal distribution of potential outcomes given covariates. We begin with the conditional quantile treatment effect:

$$\text{CQTE}(\tau \mid w) = Q_{Y_1|W}(\tau \mid w) - Q_{Y_0|W}(\tau \mid w).$$

By integrating these bounds over τ from 0 to 1, we will obtain sharp bounds for the conditional average treatment effect:

$$\text{CATE}(w) = \mathbb{E}(Y_1 \mid W = w) - \mathbb{E}(Y_0 \mid W = w).$$

Finally, averaging these bounds over the marginal distribution of W yields sharp bounds on ATE.

We first give closed form expressions for bounds on the quantile function of the potential outcome Y_x given $W = w$. Define

$$\overline{Q}^c_{Y_x|W}(\tau \mid w) = Q_{Y|X,W}\left(\min\left\{\tau + \frac{c}{p_{x|w}}\min\{\tau, 1-\tau\}, \frac{\tau}{p_{x|w}}, 1\right\} \,\Big|\, x, w\right) \tag{9}$$

and

$$\underline{Q}^c_{Y_x|W}(\tau \mid w) = Q_{Y|X,W}\left(\max\left\{\tau - \frac{c}{p_{x|w}}\min\{\tau, 1-\tau\}, \frac{\tau - 1}{p_{x|w}} + 1, 0\right\} \Bigm| x, w\right). \tag{10}$$

The following proposition and corollary formalize these results.

PROPOSITION 3—Bounds on CQTE: *Let* $w \in \text{supp}(W)$. *Let Assumptions* A1 *and* A2 *hold*. *Suppose the joint distribution of* (Y, X, W) *is known*. *Let* $\tau \in (0, 1)$. *Then* $CQTE(\tau \mid w)$ *lies in the set*

$$\begin{aligned} &\left[\underline{CQTE}^c(\tau \mid w), \overline{CQTE}^c(\tau \mid w)\right] \\ &\quad \equiv \left[\underline{Q}^c_{Y_1|W}(\tau \mid w) - \overline{Q}^c_{Y_0|W}(\tau \mid w), \overline{Q}^c_{Y_1|W}(\tau \mid w) - \underline{Q}^c_{Y_0|W}(\tau \mid w)\right]. \end{aligned} \tag{11}$$

Moreover, the interior of this set is sharp.

The bounds (11) are also sharp for the function $\text{CQTE}(\cdot \mid \cdot)$ in a sense similar to that used in Theorem 1 and Proposition 2; we omit the formal statement for brevity. This functional sharpness delivers the following result.

COROLLARY 1—Bounds on CATE and ATE: *Suppose the assumptions of Proposition* 3 *hold*. *Suppose* $\mathbb{E}(|Y| \mid X = x, W = w) < \infty$ *for all* $(x, w) \in \text{supp}(X, W)$.

1. *Then* $CATE(w)$ *lies in the set*

$$\left[\underline{CATE}^c(w), \overline{CATE}^c(w)\right] \equiv \left[\int_0^1 \underline{CQTE}^c(\tau \mid w)\, d\tau, \int_0^1 \overline{CQTE}^c(\tau \mid w)\, d\tau\right].$$

2. *Suppose further that* $\mathbb{E}[\mathbb{E}(|Y| \mid X = x, W)] < \infty$ *for* $x \in \{0, 1\}$. *Then ATE lies in the set*

$$\left[\underline{ATE}^c, \overline{ATE}^c\right] \equiv \left[\mathbb{E}\big(\underline{CATE}^c(W)\big), \mathbb{E}\big(\overline{CATE}^c(W)\big)\right]$$

assuming these means exist (*including possibly* $\pm\infty$).

Moreover, the interiors of these sets are sharp.

All of these bounds are defined directly from equations (9) and (10), or averages of those equations. Those equations have simple analytical expressions, which makes all of these bounds quite tractable. These bounds are all monotonic in c, as illustrated in Figure 2 of our numerical example. In particular, as c goes to zero, the CQTE bounds collapse to the point $Q_{Y|X,W}(\tau \mid 1, w) - Q_{Y|X,W}(\tau \mid 0, w)$ while the $\text{CATE}(w)$ bounds collapse to the point $\mathbb{E}(Y \mid X = 1, W = w) - \mathbb{E}(Y \mid X = 0, W = w)$ and the ATE bounds collapse to $\mathbb{E}[\mathbb{E}(Y \mid X = 1, W) - \mathbb{E}(Y \mid X = 0, W)]$.

Unconditional c.d.f.s and the Unconditional QTE

We can also derive bounds on the unconditional $\text{QTE}(\tau)$ by first deriving bounds on the unconditional c.d.f.s of Y_1 and Y_0 and inverting them. These unconditional c.d.f. bounds obtain by integrating the conditional bounds of Proposition 2 over w. Inverses here denote the left inverse.

COROLLARY 2—Bounds on Marginal c.d.f.s, Quantiles, and QTEs: *Let Assumptions A1 and A2 hold. Suppose the joint distribution of (Y, X, W) is known. Let $y \in \mathbb{R}$ and $\tau \in (0,1)$. Then the following hold.*

1. *$F_{Y_x}(y)$ lies in the set*

$$\left[\underline{F}^c_{Y_x}(y), \overline{F}^c_{Y_x}(y)\right] \equiv \left[\mathbb{E}\left(\underline{F}^c_{Y_x|W}(y \mid W)\right), \mathbb{E}\left(\overline{F}^c_{Y_x|W}(y \mid W)\right)\right].$$

2. *$Q_{Y_x}(\tau)$ lies in the set*

$$\left[\underline{Q}^c_{Y_x}(\tau), \overline{Q}^c_{Y_x}(\tau)\right] \equiv \left[\left(\overline{F}^c_{Y_x}\right)^{-1}(\tau), \left(\underline{F}^c_{Y_x}\right)^{-1}(\tau)\right].$$

3. *$QTE(\tau)$ lies in the set*

$$\left[\underline{QTE}^c(\tau), \overline{QTE}^c(\tau)\right] \equiv \left[\underline{Q}^c_{Y_1}(\tau) - \overline{Q}^c_{Y_0}(\tau), \overline{Q}^c_{Y_1}(\tau) - \underline{Q}^c_{Y_0}(\tau)\right].$$

Moreover, the interiors of these sets are sharp.

As with our earlier results, all three results in this corollary are also functionally sharp. And again all these bounds collapse to a single point as c approaches zero.

The ATT

The average effect of treatment on the treated is

$$\text{ATT} = \mathbb{E}(Y_1 - Y_0 \mid X = 1).$$

Under conditional independence, $\text{ATT} = \mathbb{E}[\text{CATE}(W) \mid X = 1]$. That is, we average CATE over the distribution of covariates W within the treated group, whereas ATE is the unconditional average of CATE over the covariates, $\text{ATE} = \mathbb{E}[\text{CATE}(W)]$. In Corollary 1 we showed that the bounds on ATE under conditional c-dependence are simply the average of our CATE bounds over the marginal distribution of W. Hence a natural first approach for obtaining bounds on ATT is to average our CATE bounds over the distribution of $W \mid X = 1$, just as we do in the baseline case of $c = 0$. For $c > 0$, however, this approach is not correct. This follows since, when conditional independence fails, potential outcomes are not independent of treatment assignment, even conditional on covariates. That is, for $x \in \{0, 1\}$, the distribution of $Y_x \mid X = 1, W = w$ is not necessarily the same as the distribution of $Y_x \mid X = 0, W = w$. Hence the parameters $\mathbb{E}(Y_1 - Y_0 \mid W = w)$ and $\mathbb{E}(Y_1 - Y_0 \mid X = 1, W = w)$ are not necessarily equal. Below, we derive the correct identified set for ATT under conditional c-dependence.

As in the baseline case of conditional independence, equation (1) immediately implies that $\mathbb{E}(Y_1 \mid X = 1)$ is point identified by $\mathbb{E}(Y \mid X = 1)$ without any assumptions on the dependence between Y_1 and X. Hence only $\mathbb{E}(Y_0 \mid X = 1)$ will be partially identified under deviations from conditional independence. Thus we relax Assumptions A1.3 and A2 as follows.

ASSUMPTION A1.3′: $p_1 > 0$ and, for all $w \in \text{supp}(W)$, $p_{1|w} < 1$.

ASSUMPTION A2′: X is conditionally c-dependent with Y_0 given W.

By the law of iterated expectations and some algebra,

$$\mathbb{E}(Y_0 \mid X = 1) = \frac{\mathbb{E}(Y_0) - p_0\mathbb{E}(Y \mid X = 0)}{p_1}.$$

Hence bounds on the conditional mean can be obtained from bounds on the unconditional mean $\mathbb{E}(Y_0)$. Let

$$\underline{E}_0^c(w) = \int_0^1 \underline{Q}_{Y_0}^c(\tau \mid w)\, d\tau \quad \text{and} \quad \overline{E}_0^c(w) = \int_0^1 \overline{Q}_{Y_0}^c(\tau \mid w)\, d\tau$$

denote bounds on $\mathbb{E}(Y_0 \mid W = w)$. Averaging these over the marginal distribution of W yields bounds on $\mathbb{E}(Y_0)$, denoted by

$$\underline{E}_0^c = \mathbb{E}\big(\underline{E}_0^c(W)\big) \quad \text{and} \quad \overline{E}_0^c = \mathbb{E}\big(\underline{E}_0^c(W)\big).$$

PROPOSITION 4—Bounds on ATT: *Suppose Assumptions A1.1, A1.2, A1.3′, and A2′ hold. Suppose the joint distribution of (Y, X, W) is known. Then ATT lies in the set*

$$\left[\mathbb{E}(Y \mid X = 1) - \frac{\overline{E}_0^c - p_0\mathbb{E}(Y \mid X = 0)}{p_1}, \mathbb{E}(Y \mid X = 1) - \frac{\underline{E}_0^c - p_0\mathbb{E}(Y \mid X = 0)}{p_1}\right]$$

assuming these means exist (including possibly $\pm\infty$). Moreover, the interior of this set is sharp.

Bounds for the average effect of treatment on the untreated, $\text{ATU} = \mathbb{E}(Y_1 - Y_0 \mid X = 0)$, can be obtained similarly.

Bounds With Binary Outcomes

Here we drop the continuity Assumption A1.1 and instead consider binary potential outcomes Y_x. We replace the support Assumption A1.2 by the following.

ASSUMPTION A1.2′: For all $x, x' \in \{0, 1\}$ and $w \in \text{supp}(W)$, $\text{supp}(Y_x \mid X = x', W = w) = \{0, 1\}$.

This assumption is equivalent to $\mathbb{P}(Y_x = 1 \mid X = x', W = w) \in (0, 1)$ for all $x, x' \in \{0, 1\}$ and $w \in \text{supp}(W)$. Let $p_{1|x,w} = \mathbb{P}(Y = 1 \mid X = x, W = w)$. Define

$$\overline{P}_x^c(1 \mid w) = \min\left\{\frac{p_{1|x,w}p_{x|w}}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c) + \mathbb{1}(p_{x|w} \leq c),\, p_{1|x,w}p_{x|w} + (1 - p_{x|w})\right\}$$

and

$$\underline{P}_x^c(1 \mid w) = \frac{p_{1|x,w}p_{x|w}}{\min\{p_{x|w} + c, 1\}}.$$

PROPOSITION 5: *Suppose Assumptions A1.2′, A1.3, and A2 hold. Suppose the joint distribution of (Y, X, W) is known. Let $x \in \{0, 1\}$ and $w \in \text{supp}(W)$. Then*

$$\mathbb{P}(Y_x = 1 \mid W = w) \in \big[\underline{P}_x^c(1 \mid w), \overline{P}_x^c(1 \mid w)\big].$$

Moreover, the interior of this set is sharp.

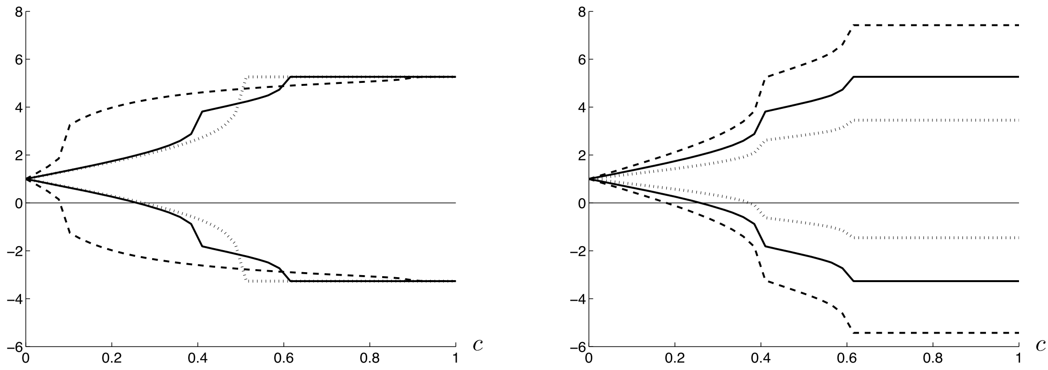

FIGURE 2.—Identified sets for ATE, and how they depend on the dgp and the value of c. Left: For three dgps, corresponding to three values of the observed propensity score $p_{1|1}$ (0.9 dashed lines, 0.6 solid lines, 0.5 dotted lines). Right: For three dgps, corresponding to three values of R^2 (15% dashed lines, 30% solid lines, 60% dotted lines).

Bounds for $\mathbb{P}(Y_x = 0 \mid W = w)$ obtain immediately by taking complements. Averaging over the marginal distribution of W yields

$$\mathbb{P}(Y_x = 1) \in \left[\mathbb{E}\big(\underline{P}_x^c(1 \mid W)\big), \mathbb{E}\big(\overline{P}_x^c(1 \mid W)\big)\right]$$

with a sharp interior.

Bounds for average treatment effects $\mathbb{P}(Y_1 = 1) - \mathbb{P}(Y_0 = 1)$ can be obtained by combining the bounds for each separate probability $\mathbb{P}(Y_x = 1)$, $x \in \{0, 1\}$, similarly to equation (11).

Numerical Illustration

We conclude this section with a brief numerical illustration. For $x \in \{0, 1\}$ and $w \in \{0, 1\}$, suppose the density of $Y \mid X = x, W = w$ is

$$f_{Y|X,W}(y \mid x, w) = \frac{1}{\gamma_X x + \gamma_W w + \sigma} \phi_{[-4,4]}\left(\frac{y - (\pi_X x + \pi_W w)}{\gamma_X x + \gamma_W w + \sigma}\right),$$

where $\phi_{[-4,4]}$ is the p.d.f. for the truncated standard normal on $[-4, 4]$. X and W are binary with

$$\mathbb{P}(X = 1) = p_1, \qquad \mathbb{P}(W = 1) = q, \quad \text{and} \quad \mathbb{P}(X = 1 \mid W = w) = p_{1|w}$$

for $w \in \{0, 1\}$. We let $(\pi_X, \pi_W) = (1, 1)$, $(\gamma_X, \gamma_W) = (0.1, 0.1)$, $p_1 = 0.5$, and $q = 0.5$ in all dgps. We specify the choice of $p_{1|w}$ and σ below.

Under the conditional independence assumption, this dgp implies that treatment effects are heterogeneous, with an average treatment effect of $\text{ATE} = \pi_X = 1$. To examine the sensitivity of this finding to partial failure of conditional independence, Figure 2 shows identified sets for ATE under conditional c-dependence for c from 0 to 1. First consider the solid lines, which are the same in both plots. These correspond to the dgp with $(p_{1|1}, p_{1|0}) = (0.6, 0.4)$ and $\sigma = 0.965$. We see that ATE under conditional independence is positive, and that this conclusion is robust to deviations of up to about $c = 0.26$ from independence, but not to larger deviations.

Next we vary the dgp parameters to examine how our identified sets depend on features of the distribution of (Y, X, W). Imbens (2003) performed similar dgp comparisons for his method using empirical data sets. In the left plot we change the observed propensity score $p_{1|w}$ while holding all other parameters fixed. Relative to the solid lines, if we increase the variation in the observed propensity score by setting $(p_{1|1}, p_{1|0}) = (0.9, 0.1)$ then the bounds widen for most values of c, as shown by the dashed lines. In particular, the conclusion that ATE is positive now only holds for c's less than about 0.085. Conversely, if we eliminate the variation in the observed propensity score by setting $(p_{1|1}, p_{1|0}) = (0.5, 0.5)$ then the bounds shrink for most values of c, as shown by the dotted lines. The conclusion that ATE is positive now holds for slightly more values of c than under the baseline dgp used for the solid lines.

Next consider the right plot. Here we change R^2 in the regression of Y on $(1, X, W)$ while holding the observed propensity score fixed at $(p_{1|1}, p_{1|0}) = (0.6, 0.4)$. We vary the value of R^2 by varying σ. The solid lines have R^2 equal to 30% (since $\sigma = 0.965$). Relative to these lines, if we decrease R^2 to 15%, then the bounds widen for all values of c, as shown by the dashed lines. The conclusion that ATE is positive becomes less robust. Conversely, if we increase R^2 to 60%, then the bounds shrink for all values of c, as shown by the dotted lines. The conclusion that ATE is positive becomes more robust.

The shape of the bounds depends on other features of the distribution of (Y, X, W) as well. For example, if π_X increases then all the identified sets shift upward. Hence, holding all else fixed, a larger ATE implies that the sign of ATE will be point identified under weaker independence assumptions. Similar analyses can also be done with other parameters of interest, like QTE(τ) for various values of τ. Here we merely illustrate the kinds of objects empirical researchers can compute using the results we develop in this paper.

4. CONCLUSION

In this paper, we studied *conditional c-dependence*, a nonparametric approach for weakening conditional independence assumptions. We used this concept to study identification of treatment effects when the conditional independence assumption partially fails, but no further data—like observations of an instrument—are available. We derived identified sets under conditional c-dependence for many parameters of interest, including average treatment effects and quantile treatment effects. These identified sets have simple, analytical characterizations. These analytical identified sets lend themselves to sample analog estimation and inference via the existing literature on inference under partial identification (see Canay and Shaikh (2017) for a survey). Our identification results can be used to analyze the sensitivity of one's results to the conditional independence assumption, without relying on auxiliary parametric assumptions.

Several questions remain. First, we focused on identification of D-parameters (Manski (2003), page 11). Many other parameters, like the variance of potential outcomes or Gini coefficients, are not D-parameters. Nonetheless, the c.d.f. and mean bounds we derived can be used as a direct input into theorem 2 of Stoye (2010) to derive explicit, analytical bounds on these spread parameters. In future work, it would be helpful to obtain precise expressions for these spread parameter bounds. Finally, while we have given several suggestions for how to interpret conditional c-dependence, there are likely other possibilities. For example, one could adapt Rosenbaum and Silber's (2009) "amplification" approach to our setting. Incorporating this or other ideas from the extensive literature on conditional treatment probabilities would be a helpful addition to our nonparametric sensitivity analysis.

APPENDIX: PROOFS

PROOF OF PROPOSITION 1: By Assumption A1.1,

$$\begin{aligned}
&\mathbb{P}(X=1 \mid Y_x=y_x, W=w) \\
&\quad = \mathbb{P}\big(X=1 \mid F_{Y_x|W}(Y_x \mid W) = F_{Y_x|W}(y_x \mid W), W=w\big) \\
&\quad = \mathbb{P}(X=1 \mid R_x=r_x, W=w),
\end{aligned}$$

where $r_x \equiv F_{Y_x|W}(y_x \mid w)$. Thus equation (2) is equivalent to equation (2′). It now suffices to show that equations (2′) and (3) are equivalent. We have

$$\begin{aligned}
&\big|f_{X,R_x|W}\big(x', r \mid w\big) - p_{x'|w} f_{R_x|W}(r \mid w)\big| \\
&\quad = \big|\mathbb{P}\big(X=x' \mid R_x=r, W=w\big) f_{R_x|W}(r \mid w) - p_{x'|w} f_{R_x|W}(r \mid w)\big| \\
&\quad = \big|\mathbb{P}\big(X=x' \mid R_x=r, W=w\big) - p_{x'|w}\big| \cdot f_{R_x|W}(r \mid w) \\
&\quad = \big|\mathbb{P}\big(X=x' \mid R_x=r, W=w\big) - p_{x'|w}\big|,
\end{aligned}$$

where the third equality follows since $R_x \mid W=w$ is uniformly distributed on [0, 1]. Hence

$$\begin{aligned}
&\sup_{r\in[0,1]} \big|\mathbb{P}\big(X=x' \mid R_x=r, W=w\big) - \mathbb{P}\big(X=x' \mid W=w\big)\big| \\
&\quad = \sup_{r\in[0,1]} \big|f_{X,R_x|W}\big(x', r \mid w\big) - p_{x'|w} f_{R_x|W}(r \mid w)\big|.
\end{aligned}$$

This holds for any $x' \in \{0, 1\}$, which completes the proof. *Q.E.D.*

The following lemma shows how to write conditional c.d.f.s as integrals of conditional probabilities. We frequently use this result below.

LEMMA 1: *Let U be a continuous random variable. Let X be a random variable with $p_x = \mathbb{P}(X=x) > 0$. Then*

$$F_{U|X}(u \mid x) = \int_{-\infty}^{u} \frac{\mathbb{P}(X=x \mid U=v)}{p_x}\, dF_U(v).$$

PROOF OF LEMMA 1: We have

$$\begin{aligned}
p_x F_{U|X}(u \mid x) &= \mathbb{P}(X=x)\mathbb{P}(U \le u \mid X=x) \\
&= \mathbb{P}(U \le u, X=x) \\
&= \mathbb{E}\big[\mathbb{1}(U \le u)\mathbb{1}(X=x)\big] \\
&= \mathbb{E}\big(\mathbb{E}\big(\mathbb{1}(U \le u)\mathbb{1}(X=x) \mid U\big)\big) \\
&= \mathbb{E}\big(\mathbb{1}(U \le u)\mathbb{E}\big(\mathbb{1}(X=x) \mid U\big)\big) \\
&= \int_{-\infty}^{\infty} \mathbb{1}(v \le u)\mathbb{E}\big(\mathbb{1}(X=x) \mid U=v\big)\, dF_U(v) \\
&= \int_{-\infty}^{u} \mathbb{P}(X=x \mid U=v)\, dF_U(v).
\end{aligned}$$

Now divide both sides by p_x. *Q.E.D.*

PROOF OF THEOREM 1: This proof has five parts: (1) Show that $\overline{F}^c_{U|X}(\cdot \mid x)$ is an upper bound. (2) Show that $\underline{F}^c_{U|X}(\cdot \mid x)$ is a lower bound. (3) Show that these bound functions are valid c.d.f.s. (4) Show that these bounds are sharp, in the sense stated in the theorem. (5) Apply these results to obtain the pointwise sharp bounds.

Part 1. We show that $F_{U|X}(u \mid x) \le \overline{F}^c_{U|X}(u \mid x)$ for all $u \in \text{supp}(U)$. Let $u \in \text{supp}(U)$ be arbitrary. First, note that

$$\begin{aligned} F_{U|X}(u \mid x) &= \int_{-\infty}^{u} \frac{\mathbb{P}(X = x \mid U = v)}{p_x} \, dF_U(v) \\ &\le \int_{-\infty}^{u} \frac{p_x + c}{p_x} \, dF_U(v) \\ &= \left(1 + \frac{c}{p_x}\right) F_U(u). \end{aligned}$$

The first line follows by Lemma 1. The second line follows by c-dependence (Assumption 2). Likewise,

$$\begin{aligned} F_{U|X}(u \mid x) &= 1 - \int_{u}^{\infty} \frac{\mathbb{P}(X = x \mid U = v)}{p_x} \, dF_U(v) \\ &\le 1 - \int_{u}^{\infty} \frac{p_x - c}{p_x} \, dF_U(v) \, dv \\ &= \left(1 - \frac{c}{p_x}\right) F_U(u) + \frac{c}{p_x}. \end{aligned}$$

Also, since

$$F_{U|X}(u \mid x) = \frac{F_U(u) - p_{1-x} F_{U|X}(u \mid 1 - x)}{p_x}$$

by the law of iterated expectations and $F_{U|X}(u \mid 1 - x) \ge 0$, we have

$$F_{U|X}(u \mid x) \le \frac{F_U(u)}{p_x}.$$

Finally, since $F_{U|X}(\cdot \mid x)$ is a c.d.f., it satisfies $F_{U|X}(u \mid x) \le 1$. Therefore, $F_{U|X}(u \mid x)$ is smaller than each of the four functions inside the minimum in the definition of $\overline{F}^c_{U|X}(u \mid x)$. Thus it is smaller than the minimum too, and hence $F_{U|X}(u \mid x) \le \overline{F}^c_{U|X}(u \mid x)$.

Part 2. We show that $F_{U|X}(u \mid x) \ge \underline{F}^c_{U|X}(u \mid x)$ for all $u \in \text{supp}(U)$. Let $u \in \text{supp}(U)$ be arbitrary. By a similar argument as in part 1,

$$\begin{aligned} F_{U|X}(u \mid x) &= \int_{-\infty}^{u} \frac{\mathbb{P}(X = x \mid U = v)}{p_x} \, dF_U(v) \\ &\ge \int_{-\infty}^{u} \frac{p_x - c}{p_x} \, dF_U(v) \\ &= \left(1 - \frac{c}{p_x}\right) F_U(u) \end{aligned}$$

and

$$\begin{aligned} F_{U|X}(u \mid x) &= 1 - \int_u^\infty \frac{\mathbb{P}(X = x \mid U = v)}{p_x}\, dF_U(v) \\ &\geq 1 - \int_u^\infty \frac{p_x + c}{p_x}\, dF_U(v) \\ &= \left(1 + \frac{c}{p_x}\right) F_U(u) - \frac{c}{p_x}. \end{aligned}$$

Also, since

$$F_{U|X}(u \mid x) = \frac{F_U(u) - p_{1-x} F_{U|X}(u \mid 1 - x)}{p_x},$$

$p_{1-x} = 1 - p_x$, and $F_{U|X}(u \mid 1 - x) \leq 1$, we have that

$$F_{U|X}(u \mid x) \geq \frac{F_U(u) - 1}{p_x} + 1.$$

Finally, since $F_{U|X}(\cdot \mid x)$ is a c.d.f., it satisfies $F_{U|X}(u \mid x) \geq 0$. Therefore, $F_{U|X}(u \mid x)$ is greater than each of the four functions inside the maximum in the definition of $\underline{F}^c_{U|X}(u \mid x)$. Thus it is greater than the maximum too, and hence $F_{U|X}(u \mid x) \geq \underline{F}^c_{U|X}(u \mid x)$.

Part 3. We show that the functions $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ are c.d.f.s on $\operatorname{supp}(U)$.

By definition, $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ are compositions of continuous functions (e.g., the function $F_U(u)$, by Assumption 2), and hence are also continuous. Also by definition, they approach 0 as u approaches $\inf \operatorname{supp}(U)$ and approach 1 as u approaches $\sup \operatorname{supp}(U)$.

When $c \leq p_x$, the expressions

$$F_U(u) - \frac{c}{p_x} \min\{F_U(u), 1 - F_U(u)\} \quad \text{and} \quad F_U(u) + \frac{c}{p_x} \min\{F_U(u), 1 - F_U(u)\}$$

are nondecreasing in u. All other arguments of $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ are also nondecreasing. Hence $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ are nondecreasing when $c \leq p_x$.

When $c > p_x$, we have

$$\begin{aligned} \overline{F}^c_{U|X}(u \mid x) &= \min\left\{F_U(u) + \frac{c}{p_x} \min\{F_U(u), 1 - F_U(u)\}, \frac{F_U(u)}{p_x}, 1\right\} \\ &= \min\left\{\left(1 + \frac{c}{p_x}\right) F_U(u), \frac{c}{p_x} + \left(1 - \frac{c}{p_x}\right) F_U(u), \frac{F_U(u)}{p_x}, 1\right\} \\ &= \min\left\{\left(1 + \frac{c}{p_x}\right) F_U(u), \frac{F_U(u)}{p_x}, 1\right\} \end{aligned}$$

since

$$\frac{c}{p_x} + \left(1 - \frac{c}{p_x}\right) F_U(u) = \frac{c}{p_x}\left[1 - F_U(u)\right] + 1 \cdot F_U(u) \geq 1$$

for all $u \in \operatorname{supp}(U)$. The inequality follows since $F_U(u) \in [0,1]$ and $c > p_x$ and therefore this term is a linear combination of 1 and something greater than 1. Each of the three terms remaining in the expression for $\overline{F}^c_{U|X}(\cdot \mid x)$ is nondecreasing in u and hence $\overline{F}^c_{U|X}(\cdot \mid x)$ is nondecreasing in u when $c > p_x$.

Likewise, when $c > p_x$,

$$\begin{aligned}\underline{F}^c_{U|X}(u \mid x) &= \max\left\{\left(1-\frac{c}{p_x}\right)F_U(u), \left(1+\frac{c}{p_x}\right)F_U(u) - \frac{c}{p_x}, \frac{F_U(u)-1}{p_x}+1, 0\right\} \\ &= \max\left\{\left(1+\frac{c}{p_x}\right)F_U(u) - \frac{c}{p_x}, \frac{F_U(u)-1}{p_x}+1, 0\right\}\end{aligned}$$

since $c > p_x$ implies

$$\left(1-\frac{c}{p_x}\right)F_U(u) \leq 0$$

for all $u \in \operatorname{supp}(U)$. Therefore, $\underline{F}^c_{U|X}(\cdot \mid x)$ is also nondecreasing when $c > p_x$. Thus we have shown that, regardless of the value of c, $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ are nondecreasing.

Putting all of these results together, we have shown that $\underline{F}^c_{U|X}(\cdot \mid x)$ and $\overline{F}^c_{U|X}(\cdot \mid x)$ satisfy all the requirements to be valid c.d.f.s.

Part 4. In this part, we prove sharpness in two steps. First, we construct a joint distribution of (U, X) consistent with Assumptions 1–4 and which yields the lower bound $\underline{F}^c_{U|X}(\cdot \mid x)$. And likewise for the upper bound $\overline{F}^c_{U|X}(\cdot \mid x)$. This yields equation (5) for $\varepsilon = 0$ and $\varepsilon = 1$. Second, we use certain linear combinations of these two joint distributions to obtain the case for $\varepsilon \in (0,1)$.

The marginal distributions of U and X are prespecified. Hence, to construct the joint distribution of (U, X), it suffices to define conditional distributions of $X \mid U$. Specifically, for $c > 0$ and $u \in \operatorname{supp}(U)$, define the conditional probabilities

$$\underline{P}^c_x(u) = \begin{cases}\max\{p_x - c, 0\} & \text{if } u \leq \underline{u}^c_x \\ \min\{p_x + c, 1\} & \text{if } \underline{u}^c_x < u,\end{cases} \quad \text{and}$$

$$\overline{P}^c_x(u) = \begin{cases}\min\{p_x + c, 1\} & \text{if } u \leq \overline{u}^c_x \\ \max\{p_x - c, 0\} & \text{if } \overline{u}^c_x < u,\end{cases}$$

where

$$\underline{u}^c_x = F_U^{-1}\left(\frac{\min\{c, 1-p_x\}}{\min\{p_x+c, 1\} - \max\{p_x - c, 0\}}\right)$$

and

$$\overline{u}^c_x = F_U^{-1}\left(1 - \frac{\min\{c, 1-p_x\}}{\min\{p_x+c, 1\} - \max\{p_x - c, 0\}}\right).$$

Note that $F_U(\overline{u}^c_x) + F_U(\underline{u}^c_x) = 1$. For $c = 0$, define $\overline{P}^c_x(u) = \underline{P}^c_x(u) = p_x$. The definition of $\underline{u}^c_x$ above is derived as the number such that the lower bound function $\underline{F}^c_{U|X}(\cdot \mid x)$ can be written as

$$\underline{F}^c_{U|X}(u \mid x) = \begin{cases} \max\left\{\left(1 - \dfrac{c}{p_x}\right)F_U(u), 0\right\} & \text{if } u \leq \underline{u}^c_x \\ \max\left\{\left(1 + \dfrac{c}{p_x}\right)F_U(u) - \dfrac{c}{p_x}, \dfrac{F_U(u)-1}{p_x} + 1\right\} & \text{if } u > \underline{u}^c_x. \end{cases}$$

An example of this two-case form is shown in Figure 1. There we also see an example of the kink point $\underline{u}^c_x$. A similar result holds for the upper bound function, using the number $\overline{u}^c_x$.

The conditional probabilities $\underline{P}^c_x(u)$ and $\overline{P}^c_x(u)$ satisfy c-dependence by construction. Moreover, they are consistent with the marginal distribution of X, $\mathbb{P}(X = x) = p_x$, since

$$\begin{aligned}
&\int_{\text{supp}(U)} \underline{P}^c_x(u)\, dF_U(u) \\
&\quad = \max\{p_x - c, 0\}\mathbb{P}\big(U \leq \underline{u}^c_x\big) + \min\{p_x + c, 1\}\mathbb{P}\big(U > \underline{u}^c_x\big) \\
&\quad = \max\{p_x - c, 0\}F_U\left(F_U^{-1}\left(\frac{\min\{c, 1-p_x\}}{\min\{p_x + c, 1\} - \max\{p_x - c, 0\}}\right)\right) \\
&\qquad + \min\{p_x + c, 1\}\left[1 - F_U\left(F_U^{-1}\left(\frac{\min\{c, 1-p_x\}}{\min\{p_x + c, 1\} - \max\{p_x - c, 0\}}\right)\right)\right] \\
&\quad = \frac{\big(\max\{p_x - c, 0\} - \min\{p_x + c, 1\}\big)\min\{c, 1-p_x\}}{\min\{p_x + c, 1\} - \max\{p_x - c, 0\}} + \min\{p_x + c, 1\} \\
&\quad = -\min\{c, 1-p_x\} + \min\{p_x + c, 1\} \\
&\quad = -\min\{c, 1-p_x\} + \min\{c, 1-p_x\} + p_x \\
&\quad = p_x
\end{aligned}$$

and, by a similar proof,

$$\int_{\text{supp}(U)} \overline{P}^c_x(u)\, dF_U(u) = p_x.$$

To see that these conditional probabilities yield our bound functions, first let $\mathbb{P}(X = x \mid U = u) = \underline{P}^c_x(u)$ for all u. Then, by Lemma 1,

$$\begin{aligned}
\mathbb{P}(U \leq u \mid X = x) &= \int_{-\infty}^{u} \frac{\underline{P}^c_x(v)}{p_x}\, dF_U(v) \\
&= \underline{F}^c_{U|X}(u \mid x).
\end{aligned}$$

To see this, note that for $u \leq \underline{u}^c_x$,

$$\begin{aligned}
\int_{-\infty}^{u} \frac{\underline{P}^c_x(v)}{p_x}\, dF_U(v) &= \frac{\max\{p_x - c, 0\}}{p_x} F_U(u) \\
&= \max\left\{\left(1 - \frac{c}{p_x}\right)F_U(u), 0\right\},
\end{aligned}$$

while for $u > \underline{u}_x^c$,

$$\begin{aligned}\int_{-\infty}^{u} \frac{\underline{P}_x^c(v)}{p_x}\, dF_U(v) &= \int_{-\infty}^{\infty} \frac{\underline{P}_x^c(v)}{p_x}\, dF_U(v) - \int_{u}^{\infty} \frac{\underline{P}_x^c(v)}{p_x}\, dF_U(v) \\ &= 1 - \min\left\{1 + \frac{c}{p_x}, \frac{1}{p_x}\right\}\bigl(1 - F_U(u)\bigr) \\ &= \max\left\{1 - \left(1 + \frac{c}{p_x}\right)\bigl(1 - F_U(u)\bigr), 1 - \frac{1 - F_U(u)}{p_x}\right\} \\ &= \max\left\{\left(1 + \frac{c}{p_x}\right)F_U(u) - \frac{c}{p_x}, \frac{F_U(u) - 1}{p_x} + 1\right\}.\end{aligned}$$

These two final expressions correspond to this lower bound. Similarly, letting $\mathbb{P}(X = x \mid U = u) = \overline{P}_x^c(u)$ for all u yields $\mathbb{P}(U \le u \mid X = x) = \overline{F}_{U|X}^c(u \mid x)$.

Thus we have shown that the bound functions are attainable. That is, equation (5) holds with $\varepsilon = 0$ or 1. Next consider $\varepsilon \in (0, 1)$. For this ε, we specify the distribution of $X \mid U$ by the conditional probability $\varepsilon \underline{P}_x^c(u) + (1 - \varepsilon)\overline{P}_x^c(u)$. This is a valid conditional probability since it is a linear combination of two terms which are between 0 and 1. Similarly, these two terms are between $p_x - c$ and $p_x + c$ and hence the linear combination is in $[p_x - c, p_x + c]$. Therefore this distribution satisfies c-dependence. By linearity of the integral and our results above, it yields

$$\int_{\text{supp}(U)} \bigl[\varepsilon \underline{P}_x^c(u) + (1 - \varepsilon)\overline{P}_x^c(u)\bigr]\, dF_U(u) = p_x$$

and hence is consistent with the marginal distribution of X. Again by linearity of the integral and our results above,

$$\mathbb{P}(U \le u \mid X = x) = \varepsilon \underline{F}_{U|X}^c(u \mid x) + (1 - \varepsilon)\overline{F}_{U|X}^c(u \mid x),$$

as needed for equation (5).

Part 5. We conclude by noting that the evaluation functional is monotonic (in the sense of first-order stochastic dominance), which yields the pointwise bounds on $F_{U|X}(u \mid x)$. These are sharp by continuity of this functional, by continuity of the equation (5) c.d.f.s in ε, and by varying ε over [0, 1]. *Q.E.D.*

PROOF OF PROPOSITION 2: First we link the observed data to the unobserved parameters of interest:

$$\begin{aligned}&F_{Y|X,W}(y \mid x, w) \\ &\quad = \mathbb{P}(Y \le y \mid X = x, W = w) \\ &\quad = \mathbb{P}(Y_x \le y \mid X = x, W = w) \\ &\quad = \mathbb{P}\bigl(F_{Y_x|W}(Y_x \mid W) \le F_{Y_x|W}(y \mid w) \mid X = x, W = w\bigr) \\ &\quad = F_{R_x|X,W}\bigl(F_{Y_x|W}(y \mid w) \mid x, w\bigr). \qquad (12)\end{aligned}$$

The second equality follows by definition (equation (1)). The third equality follows since $F_{Y_x|W}(\cdot \mid w)$ is strictly increasing (by Assumption A1.1). The fourth equality follows by definition of R_x.

The left-hand side of equation (12) is known, while the argument of the right-hand side is our parameter of interest. The main idea of this proof is that Theorem 1 yields bounds on $F_{R_x|X,W}$, which we then invert to obtain bounds on $F_{Y_x|W}$. Showing this—part (1) below—is straightforward. Several technical difficulties in proving sharpness arise, however, from the inversion step. These issues account for parts (2)–(6) below, as summarized next: (2) Define the functions $F^c_{Y_x|W}(\cdot \mid w; \varepsilon, \eta)$. (3) Show that these functions are valid c.d.f.s. (4) Show that, for any $\varepsilon \in [0,1]$, these functions can be jointly attained. (5) Show that as η converges to zero, $F_{Y_x|W}(\cdot \mid w; 0, \eta)$ approximates the lower bound from above. And likewise $F_{Y_x|W}(\cdot \mid w; 1, \eta)$ approximates the upper bound from below. (6) Show that $F_{Y_x|W}(y \mid w; \varepsilon, \eta)$ is continuous when viewed as a function of ε.

Finally, in part (7), we apply these results to obtain the pointwise bounds with sharp interior.

Part 1. First we show $F_{Y_x|W}(\cdot \mid w) \in \mathcal{F}^c_{Y_x|w}$. By

$$\mathbb{P}(X = x \mid Y_x = y, W = w) = \mathbb{P}\big(X = x \mid F_{Y_x|W}(Y_x \mid W) = F_{Y_x|W}(y \mid w), W = w\big)$$

and conditional c-dependence (Assumption A2), we have

$$\sup_{r \in [0,1]} \big|\mathbb{P}(X = x \mid R_x = r, W = w) - \mathbb{P}(X = x \mid W = w)\big| \leq c.$$

Conditioning on $W = w$, apply Theorem 1 to obtain bounds $\underline{F}^c_{R_x|X,W}$ and $\overline{F}^c_{R_x|X,W}$ for the distribution of $F_{R_x|X,W}$. Recall that $F_{R_x|W}(r \mid w) = r$ since $R_x \mid W = w \sim \text{Unif}[0,1]$.

For $\tau \in [0,1]$, let

$$\begin{aligned} \overline{Q}^c_x(\tau \mid w) &= \sup\big\{u \in [0,1] : \underline{F}^c_{R_x|X,W}(u \mid x, w) \leq \tau\big\} \\ &= \min\bigg\{\tau \frac{p_{x|w}}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c) + \mathbb{1}(p_{x|w} \leq c), \\ &\qquad \frac{p_{x|w}\tau + c}{p_{x|w} + c}, p_{x|w}\tau + (1 - p_{x|w})\bigg\} \end{aligned} \tag{13}$$

denote the *right*-inverse of $\underline{F}^c_{R_x|X,W}(\cdot \mid x, w)$. Similarly, let

$$\begin{aligned} \underline{Q}^c_x(\tau \mid w) &= \inf\big\{u \in [0,1] : \overline{F}^c_{R_x|X,W}(u \mid x, w) \geq \tau\big\} \\ &= \max\bigg\{\tau \frac{p_{x|w}}{p_{x|w} + c}, \frac{p_{x|w}\tau - c}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c), p_{x|w}\tau\bigg\} \end{aligned} \tag{14}$$

denote the *left*-inverse of $\overline{F}^c_{R_x|X,W}(\cdot \mid x, w)$.

The bounds hold trivially, by definition, for $y \notin [\underline{y}_x(w), \overline{y}_x(w))$. Suppose $y \in [\underline{y}_x(w), \overline{y}_x(w))$.

Lower bound: We have

$$\begin{aligned} \underline{F}^c_{Y_x|W}(y \mid w) &= \underline{Q}^c_x\big(F_{Y|X,W}(y \mid x, w) \mid w\big) \\ &\leq \underline{Q}^c_x\big(\overline{F}^c_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big) \mid w\big) \\ &\leq F_{Y_x|W}(y \mid w). \end{aligned}$$

The first line follows by evaluating equation (14) at $\tau = F_{Y|X,W}(y \mid x, w)$, which yields equation (7). The second line follows since, by equation (12) and since $\overline{F}^c_{R_x|X,W}$ is an upper bound,

$$F_{Y|X,W}(y \mid x, w) \leq \overline{F}^c_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big).$$

The third and final line follows by van der Vaart (2000) Lemma 21.1 part (iv), for all $y \in [\underline{y}_x(w), \overline{y}_x(w))$.

Upper bound: We similarly have

$$\begin{aligned}\overline{F}^c_{Y_x|W}(y \mid w) &= \overline{Q}^c_x\big(F_{Y|X,W}(y \mid x, w) \mid w\big) \\ &\geq \overline{Q}^c_x\big(\underline{F}^c_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big) \mid w\big) \\ &\geq F_{Y_x|W}(y \mid w).\end{aligned}$$

The first line follows by evaluating equation (13) at $\tau = F_{Y|X,W}(y \mid x, w)$, which yields equation (6). The second line follows since, by equation (12) and since $\underline{F}^c_{R_x|X,W}$ is a lower bound,

$$F_{Y|X,W}(y \mid x, w) \geq \underline{F}^c_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big).$$

The third and final line holds for all $y \in [\underline{y}_x(w), \overline{y}_x(w))$, and follows from

$$\begin{aligned}\overline{Q}^c_x\big(\underline{F}^c_{R_x|X,W}(\tau \mid x, w) \mid w\big) &= \sup\big\{u \in [0,1] : \underline{F}^c_{R_x|X,W}(u \mid x, w) \leq \underline{F}^c_{R_x|X,W}(\tau \mid x, w)\big\} \\ &\geq \tau.\end{aligned}$$

Finally, note that without support independence (Assumption A1.2) our bound functions can be too tight and thus fail to be valid bounds.

Part 2. Next we define the functions $F^c_{Y_x|W}(\cdot \mid w; \varepsilon, \eta)$. As in the proof of sharpness for Theorem 1, we will construct specific functions $\mathbb{P}(X = x \mid R_x = r, W = w)$. We then solve the equation

$$F_{Y|X,W}(y \mid x, w) = \int_0^{F_{Y_x|W}(y|w)} \frac{\mathbb{P}(X = x \mid R_x = r, W = w)}{p_{x|w}}\, dr$$

(which holds by equation (12) and Lemma 1) for $F_{Y_x|W}$ to obtain our desired functions.

For $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$ and $c > 0$, let

$$\underline{P}^c_x(r, w; \eta) = \begin{cases} \max\{p_{x|w} - c, \eta\} & \text{if } 0 \leq r \leq \underline{u}^c_x(w; \eta) \\ \min\{p_{x|w} + c, 1 - \eta\} & \text{if } \underline{u}^c_x(w; \eta) < r \leq 1 \end{cases}$$

and

$$\overline{P}^c_x(r, w; \eta) = \underline{P}^c_x(1 - r, w; \eta),$$

where

$$\underline{u}^c_x(w; \eta) = \frac{\min\{c, 1 - p_{x|w} - \eta\}}{\min\{p_{x|w} + c, 1 - \eta\} - \max\{p_{x|w} - c, \eta\}}.$$

For $c = 0$, let both of these functions be equal to $p_{x|w}$. For $\eta = 0$, these probabilities are simply those used in the proof of sharpness for Theorem 1, conditional on W. Using

$\eta < \min\{p_{x|w}, 1 - p_{x|w}\}$, it can be shown that the denominator used to define $\underline{u}_x^c(w; \eta)$ is always nonzero. This constraint on η can also be used to show that $\underline{u}_x^c(w; \eta) \in [0, 1]$. Also note that $\min\{p_{x|w}, 1 - p_{x|w}\} > 0$ by Assumption A1.3, so that such positive η's exist.

Define the function $G(\cdot; \varepsilon, \eta) : [0, 1] \to [0, 1]$ by

$$G(d; \varepsilon, \eta) = \int_0^d \frac{\varepsilon \underline{P}_x^c(r, w; \eta) + (1 - \varepsilon)\overline{P}_x^c(r, w; \eta)}{p_{x|w}} \, dr.$$

G also depends on c, x, and w but we suppress this for simplicity. For $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$ and for $\varepsilon \in [0, 1]$,

$$\varepsilon \underline{P}_x^c(r, w; \eta) + (1 - \varepsilon)\overline{P}_x^c(r, w; \eta) > 0$$

and hence $G(\cdot; \varepsilon, \eta)$ is strictly increasing; obtaining this property is a key reason why we use the η variable. $G(\cdot; \varepsilon, \eta)$ is continuous. $G(0; \varepsilon, \eta) = 0$. By derivations in part 4 below, $G(1; \varepsilon, \eta) = 1$. Thus $G(\cdot \mid \varepsilon, \eta)$ is invertible with a continuous inverse $G^{-1}(\cdot; \varepsilon, \eta)$.

We thus define $F_{Y_x|W}^c(y \mid w; \varepsilon, \eta)$ as the unique solution d^* to

$$F_{Y|X,W}(y \mid x, w) = G\big(d^*; \varepsilon, \eta\big).$$

That is,

$$F_{Y_x|W}^c(\cdot \mid w; \varepsilon, \eta) = G^{-1}\big(F_{Y|X,W}(\cdot \mid x, w); \varepsilon, \eta\big).$$

Part 3. We show that these functions are valid c.d.f.s. $F_{Y|X,W}(\cdot \mid x, w)$ is continuous and strictly increasing by Assumption A1.1. Hence, for $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$, $F_{Y_x|W}^c(\cdot \mid w; \varepsilon, \eta)$ is the composition of two continuous and strictly increasing functions, and hence itself is continuous and strictly increasing. Since $G^{-1}(0; \varepsilon, \eta) = 0$ and $G^{-1}(1; \varepsilon, \eta) = 1$, $F_{Y_x|W}^c(\cdot \mid w; \varepsilon, \eta)$ equals zero when $y = \underline{y}_x(w)$ and equals 1 when $y = \overline{y}_x(w)$. Therefore it is a valid c.d.f.

Part 4. We show that these functions can be jointly obtained. To do this, we exhibit conditional probabilities $\mathbb{P}(X = x \mid R_1 = r, W = w)$ and $\mathbb{P}(X = x \mid R_0 = r, W = w)$ such that

1. They are consistent with the c.d.f.s $F_{Y_1|W}^c(y \mid w; \varepsilon, \eta)$ and $F_{Y_0|W}^c(y \mid w; 1 - \varepsilon, \eta)$ and with the observed c.d.f. $F_{Y|X,W}$.
2. They are consistent with $p_{x|w}$.
3. They satisfy conditional c-dependence (Assumption A2) and Assumption A1.1.

Specifically, for $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$ we let

$$\mathbb{P}(X = x \mid R_1 = r, W = w) = \varepsilon \underline{P}_x^c(r, w; \eta) + (1 - \varepsilon)\overline{P}_x^c(r, w; \eta)$$

and

$$\mathbb{P}(X = x \mid R_0 = r, W = w) = (1 - \varepsilon)\underline{P}_x^c(r, w; \eta) + \varepsilon \overline{P}_x^c(r, w; \eta).$$

We now show that properties 1, 2, and 3 above hold for this choice:

1. This follows immediately by definition of the c.d.f.s $F_{Y_x|W}^c(y \mid w; \varepsilon, \eta)$ in part 2 above.

2. Recall that $R_x \mid W = w \sim \text{Unif}[0,1]$. Integrating against this marginal distribution yields

$$\begin{aligned}
&\int_0^1 \underline{P}_x^c(r, w; \eta)\,dr \\
&\quad = \max\{p_{x|w} - c, \eta\}\underline{u}_x^c(w; \eta) + \min\{p_{x|w} + c, 1-\eta\}\big(1 - \underline{u}_x^c(w; \eta)\big) \\
&\quad = \min\{p_{x|w} + c, 1-\eta\} \\
&\qquad + \frac{\min\{c, 1 - p_{x|w} - \eta\}}{\min\{p_{x|w} + c, 1-\eta\} - \max\{p_{x|w} - c, \eta\}} \\
&\qquad\quad \times \big(\max\{p_{x|w} - c, \eta\} - \min\{p_{x|w} + c, 1-\eta\}\big) \\
&\quad = \min\{p_{x|w} + c, 1-\eta\} - \min\{c, 1 - p_{x|w} - \eta\} \\
&\quad = p_{x|w} + \min\{c, 1 - p_{x|w} - \eta\} - \min\{c, 1 - p_{x|w} - \eta\} \\
&\quad = p_{x|w},
\end{aligned}$$

similarly to derivations in part 4 of the proof of Theorem 1, and

$$\begin{aligned}
\int_0^1 \overline{P}_x^c(r, w; \eta)\,dr &= \int_0^1 \underline{P}_x^c(1 - r, w; \eta)\,dr \\
&= p_{x|w}.
\end{aligned}$$

Hence

$$\int_0^1 \big[\varepsilon \underline{P}_x^c(r, w; \eta) + (1-\varepsilon)\overline{P}_x^c(r, w; \eta)\big]\,dr = p_{x|w}.$$

3. Conditional c-dependence holds by construction. Assumption A1.1 holds as shown in part 3 above.

Finally, note that we do not need to specify the joint distribution of Y_1 and Y_0 (or of R_1 and R_0) since the data only constrain the marginal distributions; any choice of copula is consistent with the data.

Part 5. We show that these functions monotonically approximate the bound functions. We first derive explicit expressions for our approximating functions. Begin with the upper bound, which corresponds to $\varepsilon = 1$. We have

$$\begin{aligned}
&F_{Y|X,W}(y \mid x, w)\,p_{x|w} \\
&\quad = \int_0^{F^c_{Y_x|W}(y|w;1,\eta)} \underline{P}_x^c(r, w; \eta)\,dr \\
&\quad = \begin{cases}
\max\{p_{x|w} - c, \eta\} F^c_{Y_x|W}(y \mid w; 1, \eta) \\
\qquad \text{if } F^c_{Y_x|W}(y \mid w; 1, \eta) \le \underline{u}_x^c(w; \eta) \\
p_{x|w} - \big[1 - F^c_{Y_x|W}(y \mid w; 1, \eta)\big]\min\{p_{x|w} + c, 1-\eta\} \\
\qquad \text{if } F^c_{Y_x|W}(y \mid w; 1, \eta) > \underline{u}_x^c(w; \eta).
\end{cases}
\end{aligned}$$

The first equality follows by definition of $F_{Y_x|W}(y \mid w; \varepsilon, \eta)$. Solving for our approximating functions yields

$$
\begin{aligned}
& F^c_{Y_x|W}(y \mid w; 1, \eta) \\
&\quad = \begin{cases} \dfrac{F_{Y|X,W}(y \mid x, w) p_{x|w}}{\max\{p_{x|w} - c, \eta\}} & \text{if this expression is} \leq \underline{u}^c_x(w; \eta) \\ \dfrac{F_{Y|X,W}(y \mid x, w) p_{x|w} + \min\{c, 1 - p_{x|w} - \eta\}}{\min\{p_{x|w} + c, 1 - \eta\}} & \text{if this expression is} > \underline{u}^c_x(w; \eta), \end{cases} \\
&\quad = \min\left\{ \frac{p_{x|w} F_{Y|X,W}(y \mid x, w)}{\max\{p_{x|w} - c, \eta\}}, \frac{F_{Y|X,W}(y \mid x, w) p_{x|w} + \min\{c, 1 - \eta - p_{x|w}\}}{\min\{p_{x|w} + c, 1 - \eta\}} \right\} \\
&\quad = \min\left\{ \frac{p_{x|w} F_{Y|X,W}(y \mid x, w)}{\max\{p_{x|w} - c, \eta\}}, \frac{F_{Y|X,W}(y \mid x, w) p_{x|w} + c}{\min\{p_{x|w} + c, 1 - \eta\}}, \right. \\
&\qquad\qquad \left. \frac{F_{Y|X,W}(y \mid x, w) p_{x|w} + (1 - \eta) - p_{x|w}}{\min\{p_{x|w} + c, 1 - \eta\}} \right\}.
\end{aligned}
$$

The last line obtains by extracting the minimum in the numerator. By similar calculations, for the lower bound ($\varepsilon = 0$) we obtain

$$
\begin{aligned}
& F^c_{Y_x|W}(y \mid w; 0, \eta) \\
&\quad = \max\left\{ \frac{p_{x|w} F_{Y|X,W}(y \mid x, w)}{p_{x|w} + c}, \right. \\
&\qquad\qquad \left. \frac{F_{Y|X,W}(y \mid x, w) p_{x|w} - \min\{c, p_{x|w} - \eta\}}{\max\{p_{x|w} - c, \eta\}}, \frac{p_{x|w} F_{Y|X,W}(y \mid x, w)}{1 - \eta} \right\}.
\end{aligned}
$$

Consider our approximation to the lower bound, $F^c_{Y_x|W}(y \mid w; 0, \eta)$. The first piece of the maximum does not depend on η, and corresponds to the first piece in the limit function $\underline{F}^c_{Y_x|W}(y \mid w)$, equation (7). For the third piece, as $\eta \searrow 0$,

$$
\frac{p_{x|w} F_{Y|X,W}(y \mid x, w)}{1 - \eta} \searrow p_{x|w} F_{Y|X,W}(y \mid x, w).
$$

This limit is the third piece of equation (7). Finally, consider the middle piece of our approximation function. If $p_{x|w} > c$ then for any $\eta \in (0, p_{x|w} - c)$, this middle piece exactly equals the middle piece of equation (7). If $p_{x|w} \leq c$ then, as $\eta \searrow 0$,

$$
\begin{aligned}
\frac{F_{Y|X,W}(y \mid x, w) p_{x|w} - \min\{c, p_{x|w} - \eta\}}{\max\{p_{x|w} - c, \eta\}} &= \frac{F_{Y|X,W}(y \mid x, w) p_{x|w} - p_{x|w} + \eta}{\eta} \\
&= \frac{\left[F_{Y|X,W}(y \mid x, w) - 1\right] p_{x|w}}{\eta} + 1 \\
&\searrow -\infty.
\end{aligned}
$$

Hence this term disappears from the overall maximum. Thus we have shown that, for any $y \in \mathbb{R}$ and $w \in \operatorname{supp}(W)$,

$$
F^c_{Y_x|W}(y \mid w; 0, \eta) \searrow \underline{F}^c_{Y_x|W}(y \mid w)
$$

as $\eta \searrow 0$. A similar argument, based on our explicit expression for $F^c_{Y_x|W}(y \mid w; 1, \eta)$, shows that

$$F^c_{Y_x|W}(y \mid w; 1, \eta) \nearrow \overline{F}^c_{Y_x|W}(y \mid w)$$

as $\eta \searrow 0$.

Part 6. We show that $F^c_{Y_x|W}(y \mid w; \cdot, \eta)$ is continuous on $[0, 1]$. Let $\{\varepsilon_n\} \subset [0, 1]$ be a sequence converging to $\varepsilon \in [0, 1]$. Recall from part 2 that, by the definition of $F^c_{Y_x|W}(y \mid w; \varepsilon, \eta)$,

$$F_{Y|X,W}(y \mid x, w) = G\big(F^c_{Y_x|W}(y \mid w; \varepsilon_n, \eta); \varepsilon_n, \eta\big)$$

for all n. Taking limits as $n \to \infty$ on both sides yields

$$F_{Y|X,W}(y \mid x, w) = G\Big(\lim_{n\to\infty} F^c_{Y_x|W}(y \mid w; \varepsilon_n, \eta); \varepsilon, \eta\Big).$$

Continuity of $G(\cdot; \cdot, \eta)$ allows us to pass the limit inside. Finally, inverting $G(\cdot; \varepsilon, \eta)$ yields

$$\begin{aligned}\lim_{n\to\infty} F^c_{Y_x|W}(y \mid w; \varepsilon_n, \eta) &= G^{-1}\big(F_{Y|X,W}(y \mid x, w); \varepsilon, \eta\big) \\ &= F^c_{Y_x|W}(y \mid w; \varepsilon, \eta),\end{aligned}$$

as desired.

Part 7. Finally, we apply these results to obtain the pointwise bounds. Fix $y \in \mathbb{R}$. Let $e \in (\underline{F}^c_{Y_x|W}(y \mid w), \overline{F}^c_{Y_x|W}(y \mid w))$. By part 5, there exists an $\eta^* > 0$ such that $e \in [F^c_{Y_x|W}(y \mid w; 0, \eta^*), F^c_{Y_x|W}(y \mid w; 1, \eta^*)]$. By part 6, $F^c_{Y_x|W}(y \mid w; \cdot, \eta^*)$ is continuous on $[0, 1]$. Hence, by the intermediate value theorem, there exists an $\varepsilon^* \in [0, 1]$ such that $e = F^c_{Y_x|W}(y \mid w; \varepsilon^*, \eta^*)$. Thus the value e is attainable. *Q.E.D.*

PROOF OF PROPOSITION 3: Recall equation (12),

$$F_{Y|X,W}(y \mid x, w) = F_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big).$$

By invertibility of $F_{Y|X,W}(\cdot \mid x, w)$,

$$y = Q_{Y|X,W}\big[F_{R_x|X,W}\big(F_{Y_x|W}(y \mid w) \mid x, w\big) \mid x, w\big].$$

By invertibility of $F_{Y_x|W}(\cdot \mid w)$, evaluate this equation at $y = Q_{Y_x|W}(\tau \mid w)$ to get

$$Q_{Y_x|W}(\tau \mid w) = Q_{Y|X,W}\big(F_{R_x|X,W}(\tau \mid x, w) \mid x, w\big).$$

Let $\underline{F}^c_{R_x|X,W}(\tau \mid x, w)$ and $\overline{F}^c_{R_x|X,W}(\tau \mid x, w)$ denote the bounds on $F_{R_x|X,W}$ obtained by applying Theorem 1 conditional on W and using $R_x \mid W = w \sim \text{Unif}[0, 1]$. This latter fact implies that these bounds are the same for R_1 and R_0. Hence we let $\underline{F}^c_{R|X,W}(\tau \mid x, w)$ and $\overline{F}^c_{R|X,W}(\tau \mid x, w)$ denote the common bounds. Substituting these bounds into the equation above yields the bounds (9) and (10). Taking the smallest and largest differences of these bounds yields (11).

Here we see that proving Proposition 3 is simpler than proving Proposition 2 since we do not need to invert the $F_{R_x|X,W}$ bounds. To complete the proof, we prove sharpness using the same construction as in that proposition. Specifically, let $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$. Then choose

$$\mathbb{P}(X = x \mid R_1 = r, W = w) = \varepsilon \underline{P}^c_x(r, w; \eta) + (1 - \varepsilon)\overline{P}^c_x(r, w; \eta)$$

and

$$\mathbb{P}(X = x \mid R_0 = r, W = w) = (1-\varepsilon)\underline{P}_x^c(r, w; \eta) + \varepsilon\overline{P}_x^c(r, w; \eta).$$

These are attainable as in part 4 of the proof of Proposition 2. By Lemma 1, we can convert these conditional probabilities to the c.d.f.s

$$\widetilde{F}^c_{R_1|X,W}(r \mid x, w; \varepsilon, \eta) = G(r; \varepsilon, \eta) \quad \text{and} \quad \widetilde{F}^c_{R_0|X,W}(r \mid x, w; \varepsilon, \eta) = G(r; 1-\varepsilon, \eta).$$

Using the properties of G, as in part 2 of the proof of Proposition 2, we see that these are valid c.d.f.s on $[0, 1]$ which are strictly increasing in r and continuous in r.

Setting $\eta = 0$ yields

$$\widetilde{F}^c_{R_x|X,W}(r \mid x, w; 1, 0) = \underline{F}^c_{R|X,W}(r \mid x, w) \quad \text{and} \quad \widetilde{F}^c_{R_x|X,W}(r \mid x, w; 0, 0) = \overline{F}^c_{R|X,W}(r \mid x, w),$$

which are not always strictly increasing. Hence, when substituted into $Q_{Y|X,W}(\cdot \mid x, w)$, the corresponding conditional quantile functions violate Assumption A1.1. As in part 5 of the proof of Proposition 2, however, we can monotonically approximate these bound functions as $\eta \searrow 0$.

Substituting our constructed c.d.f.s into our equation for $Q_{Y_x|W}(\tau \mid w)$ above and taking differences yields

$$\begin{aligned} &Q_{Y_1|W}(\tau \mid w) - Q_{Y_0|W}(\tau \mid w) \\ &\quad = Q_{Y|X,W}\big(\widetilde{F}_{R_1|X,W}(\tau \mid 1, w; \varepsilon, \eta) \mid 1, w\big) \\ &\qquad - Q_{Y|X,W}\big(\widetilde{F}_{R_0|X,W}(\tau \mid 0, w; 1-\varepsilon, \eta) \mid 0, w\big). \end{aligned} \tag{15}$$

The final step now follows as in part 7 of the proof of Proposition 2: Let $e \in (\underline{\text{CQTE}}^c(\tau \mid w), \overline{\text{CQTE}}^c(\tau \mid w))$. By monotone approximation, there exists an $\eta^* > 0$ such that e is above equation (15) evaluated at $(\varepsilon, \eta) = (1, \eta^*)$ and is below that equation evaluated at $(\varepsilon, \eta) = (0, \eta^*)$. Next note that equation (15) is continuous in ε since $Q_{Y|X,W}(\cdot \mid x, w)$ is continuous and by continuity of $\widetilde{F}_{R_x|X,W}(\tau \mid x, w; \cdot, \eta)$ on $[0, 1]$. Thus, by the intermediate value theorem, there exists an $\varepsilon^* \in [0, 1]$ such that equation (15) evaluated at $(\varepsilon, \eta) = (\varepsilon^*, \eta^*)$ yields e. *Q.E.D.*

PROOF OF COROLLARY 1:

Part 1. We obtain the CATE bounds by integrating the CQTE bounds in Proposition 3 over τ. To show sharpness, we prove two results:

(a) For any $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$,

$$\int_0^1 Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \varepsilon, \eta) \mid x, w\big)\, d\tau$$

is continuous in ε on $[0, 1]$.

(b) As $\eta \searrow 0$,

$$\int_0^1 Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) \mid x, w\big)\, d\tau \nearrow \int_0^1 Q_{Y|X,W}\big(\overline{F}^c_{R|X,W}(\tau \mid x, w) \mid x, w\big)\, d\tau$$

and

$$\int_0^1 Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 0, \eta) \mid x, w\big)\, d\tau \searrow \int_0^1 Q_{Y|X,W}\big(\underline{F}^c_{R|X,W}(\tau \mid x, w) \mid x, w\big)\, d\tau.$$

We then use the same argument as in part 7 of the proof of Proposition 2.

Proof of (a). Fix $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$. First, notice that

$$\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \varepsilon, \eta) \in \big[\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta), \widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 0, \eta)\big].$$

Hence

$$\begin{aligned} &Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \varepsilon, \eta) \mid x, w\big) \\ &\quad \in \big[Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) \mid x, w\big), Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 0, \eta) \mid x, w\big)\big], \end{aligned}$$

since $Q_{Y|X,W}(\cdot \mid x, w)$ is strictly increasing. Therefore,

$$\begin{aligned} &\big|Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \varepsilon, \eta) \mid x, w\big)\big| \\ &\quad \le \big|Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) \mid x, w\big)\big| \\ &\qquad + \big|Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 0, \eta) \mid x, w\big)\big|. \end{aligned}$$

Because of these bounds, it suffices to check that, for the endpoints $\varepsilon = 1$ and 0, the integral

$$\int_0^1 \big|Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \varepsilon, \eta) \mid x, w\big)\big|\, d\tau$$

is finite. This boundedness then allows us to use the dominated convergence theorem to pass ε limits inside the integral. Continuity of the integral then follows since $Q_{Y|X,W}(\cdot \mid x, w)$ and $\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; \cdot, \eta)$ are continuous.

We finish this part by showing that those two integrals are finite. Consider the $\varepsilon = 1$ case (the proof for $\varepsilon = 0$ is similar and is omitted). Using our definitions (part 2 of the proof of Proposition 2), computations similar to part 5 of the proof of Proposition 2 yield

$$\begin{aligned} \widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) &= \begin{cases} \max\left\{\left(1 - \dfrac{c}{p_{x|w}}\right), \dfrac{\eta}{p_{x|w}}\right\}\tau & \text{if } \tau \le \underline{u}^c_x(w; \eta) \\ 1 - \min\left\{1 + \dfrac{c}{p_{x|w}}, \dfrac{1-\eta}{p_{x|w}}\right\}(1 - \tau) & \text{if } \tau \ge \underline{u}^c_x(w; \eta) \end{cases} \\ &\equiv \begin{cases} A\tau & \text{if } \tau \le \underline{u}^c_x(w; \eta) \\ 1 - B(1 - \tau) & \text{if } \tau \ge \underline{u}^c_x(w; \eta). \end{cases} \end{aligned}$$

Hence

$$\begin{aligned} &\int_0^1 \big|Q_{Y|X,W}\big(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) \mid x, w\big)\big|\, d\tau \\ &\quad = \int_0^{\underline{u}^c_x(w;\eta)} \big|Q_{Y|X,W}(A\tau \mid x, w)\big|\, d\tau + \int_{\underline{u}^c_x(w;\eta)}^1 \big|Q_{Y|X,W}\big(1 - B(1-\tau) \mid x, w\big)\big|\, d\tau \\ &\quad = \frac{1}{A}\int_0^{A\underline{u}^c_x(w;\eta)} \big|Q_{Y|X,W}(v \mid x, w)\big|\, dv + \frac{1}{B}\int_{1-B(1-\underline{u}^c_x(w;\eta))}^1 \big|Q_{Y|X,W}(v \mid x, w)\big|\, dv \\ &\quad \le \frac{1}{A}\int_0^1 \big|Q_{Y|X,W}(v \mid x, w)\big|\, dv + \frac{1}{B}\int_0^1 \big|Q_{Y|X,W}(v \mid x, w)\big|;\, dv \end{aligned}$$

$$= \left(A^{-1} + B^{-1}\right) \cdot \mathbb{E}\left(|Y| \mid X = x, W = w\right)$$
$$< \infty.$$

The second equality follows by changing variables. The last line follows by assumption. Finally, the third line holds as follows:

- $\eta < p_{x|w}$ implies

$$A = \max\left\{\left(1 - \frac{c}{p_{x|w}}\right), \frac{\eta}{p_{x|w}}\right\} \in (0, 1]$$

and hence $0 < A\underline{u}_x^c(w; \eta) \leq \underline{u}_x^c(w; \eta) \leq 1$.

- Recalling that $\underline{u}_x^c(w; \eta) \in [0, 1]$, we have

$$1 - B\left(1 - \underline{u}_x^c(w; \eta)\right) = 1 - \min\left\{1 + \frac{c}{p_{x|w}}, \frac{1-\eta}{p_{x|w}}\right\}\left(1 - \underline{u}_x^c(w; \eta)\right) \leq 1$$

and

$$1-B\left(1-\underline{u}_x^c(w; \eta)\right) = \max\left\{\frac{c}{p_{x|w}}, \frac{p_{x|w} + \eta - 1}{p_{x|w}}\right\} + \underline{u}_x^c(w; \eta) \min\left\{\left(1+\frac{c}{p_{x|w}}\right), \frac{1-\eta}{p_{x|w}}\right\} \geq 0.$$

Finally, note that $B > 0$.

Proof of (b). From the proof of (a),

$$\int_0^1 Q_{Y|X,W}\left(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 1, \eta) \mid x, w\right) d\tau < \infty$$

and

$$\int_0^1 Q_{Y|X,W}\left(\widetilde{F}^c_{R_x|X,W}(\tau \mid x, w; 0, \eta) \mid x, w\right) d\tau > -\infty$$

for any $\eta \in (0, \min\{p_{x|w}, 1 - p_{x|w}\})$. The integrands converge pointwise monotonically to the limit functions $\overline{F}^c_{R|X,W}(\tau \mid x, w)$ and $\underline{F}^c_{R|X,W}(\tau \mid x, w)$, respectively, as $\eta \searrow 0$. The result now follows by the monotone convergence theorem (e.g., Theorem 4.3.2 on page 131 of Dudley (2002)).

Part 2. The ATE bounds follow by integrating the CATE bounds over the marginal distribution of W. Next we show sharpness. We consider the case $c > 0$, so that our bounds have nonempty interior. The case $c = 0$ is trivial as this is the usual unconfoundedness result.

Let

$$C(w) = \mathbb{E}(Y \mid X = 1, W = w) - \mathbb{E}(Y \mid X = 0, W = w).$$

$C(w)$ is in the identified set for CATE(w). Moreover, $\mathbb{E}(|C(W)|) < \infty$ by assumption. Define the functional H on the set of functions f satisfying $\mathbb{E}(|f(W)|) < \infty$ by

$$H(f) = \int_{\text{supp}(W)} f(w)\, dF_W(w).$$

H is continuous in the sup-norm, monotonic in the pointwise order on functions, and ATE $= H(\text{CATE}(\cdot))$, assuming this mean exists. $H(C)$ is finite. Since $c > 0$, $H(C) \in (\underline{\text{ATE}}^c, \overline{\text{ATE}}^c)$.

Let $e \in (\underline{\text{ATE}}^c, \overline{\text{ATE}}^c)$. We want to find a function $f(\cdot)$ such that

$$f(w) \in \big(\underline{\text{CATE}}^c(w), \overline{\text{CATE}}^c(w)\big)$$

and $H(f) = e$. First suppose $e - H(C) > 0$. Then we are looking for a function f that is sufficiently above C, but does not violate the CATE bounds. Define

$$f_M(w) = \min\big\{\overline{\text{CATE}}^c(w), C(w) + M\big\}$$

for $M > 0$. Then $H(f_0) = H(C)$. Moreover, for each $w \in \text{supp}(W)$, $f_M(w) \nearrow \overline{\text{CATE}}^c(w)$ as $M \nearrow \infty$. Thus $H(f_M) \nearrow H(\overline{\text{CATE}}^c(\cdot)) = \overline{\text{ATE}}^c$ by the monotone convergence theorem. Thus, since $e < \overline{\text{ATE}}^c$, there exists an $\overline{M}$ such that $H(f_{\overline{M}}) > e$. Finally, we note that $f_M(w)$ is continuous in M and hence $H(f_M)$ is continuous in M. Thus, by the intermediate value theorem, there exists an $0 < M^* < \overline{M}$ such that $H(f_{M^*}) = e$. Thus e is attainable. A similar argument applies if $e - H(C) \leq 0$. *Q.E.D.*

PROOF OF COROLLARY 2:

Part 1. We obtain bounds on F_{Y_x} by integrating the bounds in Proposition 2 with respect to the marginal distribution of W. For $\delta \in (0, 1)$, let $\eta(w, \delta) = \delta \cdot \min\{p_{x|w}, 1 - p_{x|w}\}$. Note that $0 < \eta(w, \delta) < \min\{p_{x|w}, 1 - p_{x|w}\}$. By the dominated convergence theorem,

$$\begin{aligned}\lim_{\delta \searrow 0} \mathbb{E}\big[F^c_{Y_x|W}(y \mid W; 0, \eta(W, \delta))\big] &= \mathbb{E}\Big(\lim_{\delta \searrow 0} F^c_{Y_x|W}(y \mid W; 0, \eta(W, \delta))\Big) \\ &= \underline{F}_{Y_x}(y).\end{aligned}$$

The second line follows by Proposition 2. A similar argument applies for the upper bound. Also by the dominated convergence theorem, $\mathbb{E}[F^c_{Y_x|W}(y \mid W; \varepsilon, \eta(W, \delta))]$ is continuous in $\varepsilon \in [0, 1]$, for all $\delta \in (0, 1)$. The argument now proceeds as in part 7 of the proof of Proposition 2.

Part 2. The bounds on $Q_{Y_x}(\tau)$ are just the inverse of the bounds on $F_{Y_x}(\tau)$ from part 1. Let $\eta(w, \delta)$ be defined as in part 1. As $\delta \searrow 0$, the c.d.f. $\mathbb{E}[F^c_{Y_x|W}(y \mid W; 0, \eta(W, \delta))]$ converges pointwise to $\underline{F}_{Y_x}(y)$ from above by part 1. Therefore, for any sequence $1 > \delta_n \geq \delta_{n+1} \geq \cdots > 0$,

$$\begin{aligned}&\inf\big\{y \in \mathbb{R} : \mathbb{E}\big[F^c_{Y_x|W}(y \mid W; 0, \eta(W, \delta_n))\big] \geq \tau\big\} \\ &\quad \leq \inf\big\{y \in \mathbb{R} : \mathbb{E}\big[F^c_{Y_x|W}(y \mid W; 0, \eta(W, \delta_{n+1}))\big] \geq \tau\big\} \\ &\quad \leq \cdots.\end{aligned}$$

This sequence thus converges monotonically to

$$\inf\big\{y \in \mathbb{R} : \underline{F}_{Y_x}(y) \geq \tau\big\} = \overline{Q}_{Y_x}(\tau),$$

which may be $+\infty$. A similar argument applies for the lower bound. Moreover, for any fixed $\delta \in (0, 1)$, the inverse of $\mathbb{E}[F^c_{Y_x|W}(\cdot \mid W; \varepsilon, \eta(W, \delta))]$ at τ is continuous in ε over $[0, 1]$ (by a proof similar to part 6 of the proof of Proposition 2). The argument now proceeds as in part 7 of the proof of Proposition 2.

Result 3. This result for $\text{QTE}(\tau)$ follows by combining the bounds on $Q_{Y_1}(\tau)$ and $Q_{Y_0}(\tau)$ from part 2, analogously to equation (11), noting that the joint identified set is the product of the marginal identified sets. *Q.E.D.*

PROOF OF PROPOSITION 4: By the law of total probability and the definition of ATT,

$$\text{ATT} = \mathbb{E}(Y \mid X=1) - \frac{\mathbb{E}(Y_0) - p_0 \mathbb{E}(Y \mid X=0)}{p_1}. \tag{16}$$

Using Assumptions A1.1, A1.2, A1.3′, and A2′, a proof similar to that of Corollary 1 shows that the set $(\underline{E}_0^c, \overline{E}_0^c)$ is sharp. Substituting these bounds in equation (16) completes the proof. *Q.E.D.*

PROOF OF PROPOSITION 5: First, we show that $[\underline{P}_x^c(1 \mid w), \overline{P}_x^c(1 \mid w)]$ are bounds for $\mathbb{P}(Y_x = 1 \mid W = w)$. We then show sharpness of the interior.

When $c < p_{x|w}$,

$$\begin{aligned}
&\mathbb{P}(Y_x = 1 \mid W = w) \\
&\quad = \frac{\mathbb{P}(Y_x = 1, W = w)}{\mathbb{P}(W = w)} \frac{\mathbb{P}(Y_x = 1, X = x, W = w)}{\mathbb{P}(Y_x = 1, X = x, W = w)} \frac{\mathbb{P}(X = x, W = w)}{\mathbb{P}(X = x, W = w)} \\
&\quad = \frac{p_{1|x,w} p_{x|w}}{\mathbb{P}(X = x \mid Y_x = 1, W = w)} \\
&\quad \leq \frac{p_{1|x,w} p_{x|w}}{p_{x|w} - c}.
\end{aligned}$$

The last line follows by conditional c-dependence. Also,

$$\begin{aligned}
&\mathbb{P}(Y_x = 1 \mid W = w) \\
&\quad = \mathbb{P}(Y = 1 \mid X = x, W = w) p_{x|w} + \mathbb{P}(Y_x = 1 \mid X = 1 - x, W = w)(1 - p_{x|w}) \\
&\quad \leq p_{1|x,w} p_{x|w} + 1 \cdot (1 - p_{x|w}).
\end{aligned}$$

Combining these two inequalities yields $\mathbb{P}(Y_x = 1 \mid W = w) \leq \overline{P}_x^c(1 \mid w)$.

Similarly,

$$\begin{aligned}
\mathbb{P}(Y_x = 1 \mid W = w) &= \frac{p_{1|x,w} p_{x|w}}{\mathbb{P}(X = x \mid Y_x = 1, W = w)} \\
&\geq \frac{p_{1|x,w} p_{x|w}}{\min\{p_{x|w} + c, 1\}}
\end{aligned}$$

by conditional c-dependence and by $\mathbb{P}(X = x \mid Y_x = 1, W = w) \leq 1$.

To show sharpness of the interior, fix $p^* \in (\underline{P}_x^c(1 \mid w), \overline{P}_x^c(1 \mid w))$. We want to exhibit a joint distribution of (Y_x, X, W) consistent with the data, our assumptions, and which yields this element p^*. Since the distribution of (X, W) is observed, we only need to specify the distribution of $Y_x \mid X, W$. Since Y_x and X are binary, there are only two parameters to this distribution to specify, for each $w \in \text{supp}(W)$. The first is $\mathbb{P}(Y_x = 1 \mid X = x, W = w) = \mathbb{P}(Y = 1 \mid X = x, W = w)$, which is point identified from the data. Hence the only unknown parameter is the value $\mathbb{P}(Y_x = 1 \mid X = 1 - x, W = w)$, which must be chosen such that

1. $\mathbb{P}(Y_x = 1 \mid W = w) = p^*$,
2. $\mathbb{P}(Y_x = 1 \mid X = 1 - x, W = w) \in (0, 1)$, and
3. conditional c-dependence is satisfied.

Proof of 1. Choose

$$\mathbb{P}(Y_x = 1 \mid X = 1-x, W = w) = \frac{p^* - p_{1|x,w}p_{x|w}}{1 - p_{x|w}}.$$

Then,

$$\begin{aligned}
&\mathbb{P}(Y_x = 1 \mid W = w) \\
&\quad = \mathbb{P}(Y = 1 \mid X = x, W = w)p_{x|w} + \mathbb{P}(Y_x \mid X = 1-x, W = w)(1 - p_{x|w}) \\
&\quad = p^*.
\end{aligned}$$

Proof of 2. We have

$$\begin{aligned}
\mathbb{P}(Y_x = 1 \mid X = 1-x, W = w) &= \frac{p^* - p_{1|x,w}p_{x|w}}{1 - p_{x|w}} \\
&> \frac{\underline{P}_x^c(1 \mid w) - p_{1|x,w}p_{x|w}}{1 - p_{x|w}} \\
&= \frac{p_{1|x,w}p_{x|w}}{1 - p_{x|w}}\left(\frac{1}{\min\{p_{x|w} + c, 1\}} - 1\right) \\
&\geq 0,
\end{aligned}$$

where the second line follows by our choice of p^*. Similarly,

$$\begin{aligned}
&\frac{p^* - p_{1|x,w}p_{x|w}}{1 - p_{x|w}} \\
&\quad < \frac{\overline{P}_x^c(1 \mid w) - p_{1|x,w}p_{x|w}}{1 - p_{x|w}} \\
&\quad = \frac{\min\left\{\frac{p_{1|x,w}p_{x|w}}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c) + \mathbb{1}(p_{x|w} \leq c),\ p_{1|x,w}p_{x|w} + (1 - p_{x|w})\right\} - p_{1|x,w}p_{x|w}}{1 - p_{x|w}} \\
&\quad = \min\left\{\frac{1}{1 - p_{x|w}}\left[\frac{p_{1|x,w}p_{x|w}}{p_{x|w} - c}\mathbb{1}(p_{x|w} > c) + \mathbb{1}(p_{x|w} \leq c) - p_{1|x,w}p_{x|w}\right], 1\right\} \\
&\quad \leq 1.
\end{aligned}$$

Hence $\mathbb{P}(Y_x = 1 \mid X = 1-x, W = w) \in (0,1)$ is a valid probability which satisfies Assumption A1.2′.

Proof of 3. By Bayes' rule, we have

$$\mathbb{P}(X = 1 \mid Y_x = 1, W = w) = \frac{p_{1|1,w}p_{1|w}}{\mathbb{P}(Y_x = 1 \mid W = w)}.$$

By the lower bound on $\mathbb{P}(Y_x = 1 \mid W = w)$, we have

$$\mathbb{P}(X = 1 \mid Y_x = 1, W = w) \leq \frac{p_{1|1,w}p_{1|w}}{\underline{P}_1^c(1 \mid w)} = \min\{p_{1|w} + c, 1\} \leq p_{1|w} + c.$$

By the upper bound on $\mathbb{P}(Y_x = 1 \mid W = w)$, we have

$$
\begin{aligned}
&\mathbb{P}(X = 1 \mid Y_x = 1, W = w) \\
&\quad \geq \frac{p_{1|1,w} p_{1|w}}{\overline{P}_1^c(1 \mid w)} \\
&\quad = \frac{p_{1|1,w} p_{1|w}}{\min\left\{ \frac{p_{1|1,w} p_{1|w}}{p_{1|w} - c} \mathbb{1}(p_{1|w} > c) + \mathbb{1}(p_{1|w} \leq c), \, p_{1|1,w} p_{1|w} + (1 - p_{1|w}) \right\}}.
\end{aligned}
$$

When $p_{1|w} > c$, this gives

$$
\begin{aligned}
\mathbb{P}(X = 1 \mid Y_x = 1, W = w) &\geq \max\left\{ p_{1|w} - c, \, \frac{p_{1|1,w} p_{1|w}}{p_{1|1,w} p_{1|w} + (1 - p_{1|w})} \right\} \\
&\geq p_{1|w} - c.
\end{aligned}
$$

If $p_{1|w} \leq c$, then $0 \geq p_{1|w} - c$ and hence $\mathbb{P}(X = 1 \mid Y_x = 1, W = w) \geq p_{1|w} - c$ holds trivially.

Therefore $\mathbb{P}(X = 1 \mid Y_x = 1, W = w) \in [p_{1|w} - c, \, p_{1|w} + c]$. A similar calculation yields the same result for $\mathbb{P}(X = 1 \mid Y_x = 0, W = w)$. *Q.E.D.*

REFERENCES

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113, 151–184. [318,321]

——— (2008): "Using Selection on Observed Variables to Assess Bias From Unobservables When Evaluating Swan–Ganz Catheterization," *American Economic Review P&P*, 98, 345–350. [318,321]

CANAY, I. A., AND A. M. SHAIKH (2017): "Practical and Theoretical Advances in Inference for Partially Identified Models," in *Advances in Economics and Econometrics: 11th World Congress of the Econometric Society (Forthcoming)*. Cambridge University Press. [331]

CORNFIELD, J., W. HAENSZEL, E. C. HAMMOND, A. M. LILIENFELD, M. B. SHIMKIN, AND E. L. WYNDER (1959): "Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions," *Journal of the National Cancer Institute*, 22, 173–203. [321,322]

CROSS, P. J., AND C. F. MANSKI (2002): "Regressions, Short and Long," *Econometrica*, 70, 357–368. [324]

DUDLEY, R. M. (2002): *Real Analysis and Probability*. Cambridge University Press. [346]

HOTZ, V. J., C. H. MULLIN, AND S. G. SANDERS (1997): "Bounding Causal Effects Using Data From a Contaminated Natural Experiment: Analysing the Effects of Teenage Childbearing," *The Review of Economic Studies*, 64, 575–603. [319]

ICHINO, A., F. MEALLI, AND T. NANNICINI (2008): "From Temporary Help Jobs to Permanent Employment: What can We Learn From Matching Estimators and Their Sensitivity?" *Journal of Applied Econometrics*, 23, 305–327. [318]

IMBENS, G. W. (2000): "The Role of the Propensity Score in Estimating Dose-Response Functions," *Biometrika*, 87, 706–710. [319]

——— (2003): "Sensitivity to Exogeneity Assumptions in Program Evaluation," *American Economic Review P&P*, 93 126–132. [318,321,322,331]

IMBENS, G. W., AND D. B. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press. [317,318]

LIN, D. Y., B. M. PSATY, AND R. A. KRONMAL (1998): "Assessing the Sensitivity of Regression Results to Unmeasured Confounders in Observational Studies," *Biometrics*, 54 948–963. [318]

MANSKI, C. F. (1989): "Anatomy of the Selection Problem," *Journal of Human Resources*, 24, 343–360. [319]

——— (1990): "Nonparametric Bounds on Treatment Effects," *American Economic Review P&P*, 80, 319–323. [319]

——— (2003): *Partial Identification of Probability Distributions*. Springer. [331]

——— (2007): *Identification for Prediction and Decision*. Harvard University Press. [318,324]

——— (2016): "Credible Interval Estimates for Official Statistics With Survey Nonresponse," *Journal of Econometrics*, 191, 293–301. [318]

MANSKI, C. F., AND J. V. PEPPER (2000): "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68, 997–1010. [319]

——— (2009): "More on Monotone Instrumental Variables," *The Econometrics Journal*, 12, S200–S216. [319]

MASTEN, M. A., AND A. POIRIER (2016): "Partial Independence in Nonseparable Models," Cemmap Working Paper, CWP26/16. [317,318]

——— (2017): "Inference on Breakdown Frontiers," Cemmap Working Paper, CWP20/17. [317]

MOLINARI, F., AND M. PESKI (2006): "Generalization of a Result on "Regressions, Short and Long"," *Econometric Theory*, 22, 159–163. [324]

OSTER, E. (2018): "Unobservable Selection and Coefficient Stability: Theory and Evidence," Journal of Business & Economic Statistics (forthcoming). [321]

ROBINS, J. M., A. ROTNITZKY, AND D. O. SCHARFSTEIN (2000): "Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models," in *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. Springer, 1–94. [318,321]

ROSENBAUM, P. R. (1995): *Observational Studies*. Springer-Verlag, New York. [318]

——— (2002a): *Observational Studies* (second Ed.). Springer-Verlag, New York. [318]

——— (2002b): "Rejoinder to "Covariance Adjustment in Randomized Experiments and Observational Studies"," *Statistical Science*, 17, 321–327. [322]

ROSENBAUM, P. R., AND D. B. RUBIN (1983): "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study With Binary Outcome," *Journal of the Royal Statistical Society, Series B*, 45 212–218. [318,321,322]

ROSENBAUM, P. R., AND J. H. SILBER (2009): "Amplification of Sensitivity Analysis in Matched Observational Studies," *Journal of the American Statistical Association*, 104, 1398–1405. [331]

STOYE, J. (2010): "Partial Identification of Spread Parameters," *Quantitative Economics*, 1, 323–357. [331]

VAN DER VAART, A. W. (2000): *Asymptotic Statistics*. Cambridge University Press. [339]

Co-editor Elie Tamer handled this manuscript.

Manuscript received 17 June, 2016; final version accepted 18 September, 2017; available online 19 September, 2017.