

A General Approach to Relaxing Unconfoundedness*

Matthew A. Masten[†] Alexandre Poirier[‡] Muyang Ren[§]

January 26, 2025

Abstract

This paper defines a general class of relaxations of the unconfoundedness assumption. This class includes several previous approaches as special cases, including the marginal sensitivity model of Tan (2006). This class therefore allows us to precisely compare and contrast these previously disparate relaxations. We use this class to derive a variety of new identification results which can be used to assess sensitivity to unconfoundedness. In particular, the prior literature focuses on average parameters, like the average treatment effect (ATE). We move beyond averages by providing sharp bounds for a large class of parameters, including both the quantile treatment effect (QTE) and the distribution of treatment effects (DTE), results which were previously unknown even for the marginal sensitivity model.

JEL classification: C14; C18; C21; C25; C51

Keywords: Identification, Treatment Effects, Partial Identification, Sensitivity Analysis, Unconfoundedness

*We thank participants at the 2024 Southern Economic Association annual meeting. Masten thanks the National Science Foundation for research support under Grant 1943138.

[†]Department of Economics, Duke University, matt.masten@duke.edu

[‡]Department of Economics, Georgetown University, alexandre.poirier@georgetown.edu

[§]Department of Economics, Duke University, muyang.ren@duke.edu

1 Introduction

A large literature studies the identification and estimation of treatment effects when a binary treatment is randomly assigned conditional on covariates. This assumption is called unconfoundedness, conditionally independent treatment assignment, or ignorability, among other terms. With observational data it is often considered very strong, however, so a corresponding literature has developed to relax this assumption. These papers use a variety of different classes of relaxations of unconfoundedness. That is, there are different ways of formalizing the idea that treatment is “almost” randomly assigned, given the covariates. This variation raises a question: How do these different relaxations compare to each other? This question is important because empirical researchers are often concerned that the number of robustness checks they must consider is constantly growing; if some of these checks are related, however, then that relationship can potentially be used to simplify the overall analysis. Moreover, mathematically related analyses do not necessarily provide “independent” evidence of robustness, a second motivation for better understanding the relationships between different relaxations of an assumption.

With that aim, this paper makes two main contributions. First, we define a general class of relaxations, which includes several previous approaches as special cases. Second, we derive closed form, analytical identification results for treatment effects under this general class of relaxations. This paper therefore unifies several disparate identification results in the literature. In doing so, we also provide a variety of new identification results, because we study an extensive list of parameters, including quantile treatment effects (QTEs) and the distribution of treatment effects (DTEs), whereas most existing papers focus solely on average-type treatment effects. These new results were previously unknown even for the specific types of relaxations that have been considered before. We give a precise discussion of how our results compare to the previous literature in the next subsection.

In section 2 we set up the baseline treatment effects model and define the target parameters we study. We define our general class of relaxations at the start of section 3. We show how this class relates to previous relaxations in sections 3.1 and 3.2. In section 4 we derive general analytical identification results for marginal cdfs of potential outcomes and monotonic functionals of those cdfs. We apply those results in section 5 to obtain analytical bounds on various treatment effect parameters. We conclude in section 6.

Related Literature

A vast literature studies unconfoundedness; we do not attempt a comprehensive review here. Instead we discuss the most closely related prior work. Nonparametric relaxations of unconfoundedness were pioneered by Paul Rosenbaum’s work (see his 2002 or 2017 books for a survey, for example). His work focuses on sensitivity analysis within the context of finite sample randomization inference (c.f., chapter 5 of Imbens and Rubin 2015). Much of the subsequent literature has instead focused on large population level identification analysis. In particular, inspired by Rosenbaum’s approach, Tan (2006) proposed the *marginal sensitivity model* (MSM), a specific nonparametric relaxation of unconfoundedness (which we review in section 3). Given this relaxation, Tan showed that bounds on parameters of interest can be characterized as the solutions to optimization problems with infinitely many constraints, but did not provide any formal results, proofs, or closed form expressions for these bounds. Zhao, Small, and Bhattacharya (2019) derived non-sharp bounds on the average potential outcome $\mathbb{E}(Y_x)$ and the average treatment effect (ATE) under the MSM, but also did not derive closed form expressions for these bounds. Dorn and Guo (2023) strengthened that result by deriving sharp bounds on $\mathbb{E}(Y_x)$, ATE, and the average effect of treatment on the treated (ATT) under the MSM, but again without closed form expressions. Dorn, Guo, and Kallus (2024) subsequently refined that result by obtaining closed form expressions for sharp bounds on $\mathbb{E}(Y_x)$ and ATE under the MSM, in addition to developing the concept of double-validity and double sharpness. Tan (2024) gives alternative sharp bound expressions for the ATE under the MSM. Kallus and Zhou (2018) studied policy learning under the MSM, which is related to identification of the average weighted welfare (what they call the “policy value”), but they do not derive population bounds on this parameter.

This existing literature on the MSM largely focuses on average potential outcomes $\mathbb{E}(Y_x)$ or the ATE. Our paper provides the first sharp bounds on a wide variety of target parameters under the MSM, including the quantile treatment effect (QTE), the quantile treatment effect on the treated (QTT), the distribution of treatment effects (DTE), and the average weighted welfare (AWW). Moreover, for many of the parameters we study, our bounds are closed form. The existence of closed form expressions simplifies the construction of estimation and inference procedures, and also allows us to analytically examine how the bounds depend on the distribution of the observed data, and thus which features of the data lead results to be robust.

Masten and Poirier (2016, 2018a) proposed an alternative relaxation of unconfoundedness called *conditional c-dependence*, and derived closed form sharp bounds on a variety of treatment effect parameters under this relaxation, including $\mathbb{E}(Y_x)$, ATE, ATT, the QTE,

and the DTE (in Masten and Poirier 2020). In the current paper we extend these identification results to a class of parameters that also includes the average weighted welfare (AWW), weighted average treatment effects, and to quantiles of the distribution of conditional average treatment effects (QCATE). Those earlier papers also restricted attention to continuous or binary outcomes whereas our new results apply for any distribution of the outcome, including mixed continuous-discrete distributions. We also show how the conditional c -dependence relaxation is related to the marginal sensitivity model.

While our general class of relaxations includes several previously proposed relaxations of unconfoundedness, there are alternative relaxations where it is not yet clear if they can be accommodated by our class. This includes Bonvini and Kennedy (2022), who derive closed-form sharp bounds on ATE under a mixture-style relaxation, and Huang and Pimentel (2025), who derive closed-form non-sharp bounds on ATT under an assumption about how much unobserved variables can affect the variance of odds ratios similar to those that arise in the MSM; in appendix A.6 they derive sharp but non-closed form bounds on the ATT under the same relaxation. It also includes Ding and VanderWeele (2016) and VanderWeele and Ding (2017), who derive closed-form non-sharp bounds on the causal relative risk under assumptions about relative risks involving latent confounders; Sjölander (2024) derives closed-form sharp bounds under the same relaxation.

There are several related papers that provide general methods for deriving bounds. Dorn and Yap (2024) show how to derive analytical expressions for sharp bounds on parameters that can be written as certain weighted averages of outcomes under a restriction on a generalized likelihood ratio. Like us, they show that their class of relaxations includes several previous relaxations (such as the MSM of Tan 2006 and conditional c -dependence of Masten and Poirier 2018a). Whereas we only study relaxations of unconfoundedness, they also show how to use their results to do sensitivity analysis for instrumental variables and regression discontinuity designs. Their analysis of unconfoundedness, however, focuses on average potential outcomes and ATE, whereas we also study parameters like the QTE and DTE. A large literature in econometrics has studied how to derive identified sets for a variety of parameters under a variety of assumptions when all observed variables are discretely distributed; see, for example, Torgovitsky (2019) and Gu, Russell, and Stringham (2024), and the references therein. Duarte (2024) uses similar ideas to numerically compute identified sets for a variety of sensitivity analyses when all variables are discretely distributed. Rambachan, Coston, and Kennedy (2023) provide a variety of general sensitivity analyses for binary outcomes. In contrast to this literature, we obtain analytical sharp bounds without any restriction on the distribution of the outcome variable, which allows the outcome to be continuously distributed or even mixed continuous-discretely distributed.

Several prior papers also discuss the relationship between various relaxations of unconfoundedness. Masten and Poirier (2023) discuss mean independence, quantile independence assumptions, and a weaker version of quantile independence that they call \mathcal{U} -independence. Their focus is on interpreting relaxations in terms of treatment selection models, rather than providing identification results for a broad class of relaxations. Zhao et al. (2019, section 7.2) discuss the relationship between the MSM and Rosenbaum’s sensitivity model. For binary outcomes, Rambachan et al. (2023, appendix D) relate their relaxation to the MSM, Rosenbaum’s sensitivity model, conditional c -dependence, and an approach called Tukey’s factorization.

Notation

For random a variable A and a random vector B , we let $F_{A|B}(a | b) := \mathbb{P}(A \leq a | B = b)$ denote the conditional cdf. For $\tau \in (0, 1)$, we let $Q_{A|B}(\tau | b) := \inf\{a \in \mathbb{R} : F_{A|B}(a | b) \geq \tau\}$ denote the left-inverse of this cdf, that is, its conditional quantile function.

2 Setup and Target Parameters

We are interested in the causal impact of a binary treatment variable $X \in \{0, 1\}$ on an outcome variable Y . Let (Y_1, Y_0) be potential outcomes under treatment and no treatment respectively. Denote the realized outcome by

$$Y = XY_1 + (1 - X)Y_0.$$

Let W be a vector of covariates with support $\text{supp}(W) \subseteq \mathbb{R}^{d_w}$. We use $p_{x|w}$ to denote $\mathbb{P}(X = x | W = w)$. $p_{1|w}$ thus denotes the propensity score. We assume realizations of (Y, X, W) are observed by the researcher. Our identification analysis abstracts from sampling uncertainty and assumes the joint distribution of (Y, X, W) is known.

Throughout the paper we maintain the following assumption, which is a strict overlap assumption. It is also sometimes called strict positivity.

Assumption 1. There exists $\epsilon > 0$ such that $p_{1|w} \in [\epsilon, 1 - \epsilon]$ for all $w \in \text{supp}(W)$.

With observational data, a commonly imposed assumption is unconfoundedness. It is also called selection on observables, ignorability, or conditional independence. This assumption states that potential outcomes are independent of treatment given covariates W . This conditional independence is either imposed jointly on both potential outcomes, or on each

potential outcome separately:

$$Y_x \perp\!\!\!\perp X \mid W \text{ for } x \in \{0, 1\} \quad \text{or} \quad (Y_1, Y_0) \perp\!\!\!\perp X \mid W. \quad (1)$$

Under Assumption 1 and either version of unconfoundedness given in equation (1), it is well known that the pair of distribution functions $(F_{Y_1|W}, F_{Y_0|W})$ is point-identified via

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

for $x \in \{0, 1\}$. This point-identification implies that many parameters that summarize aspects of the distribution of (Y_1, Y_0, X, W) are point-identified. Specifically, parameters that can be expressed as functionals of $F_{Y_1|W}$, $F_{Y_0|W}$, and the known distribution of observables (Y, X, W) , are point-identified. For example, we can point-identify the Conditional Average Treatment Effect (CATE) as defined by $\text{CATE}(w) := \mathbb{E}[Y_1 - Y_0 \mid W = w]$ because it can be written as

$$\mathbb{E}[Y_1 - Y_0 \mid W = w] = \int y_1 dF_{Y_1|W}(y_1 \mid w) - \int y_0 dF_{Y_0|W}(y_0 \mid w).$$

However, parameters that depend on other aspects of the distribution of potential outcomes may only be partially identified. For example, consider the distribution function of $Y_1 - Y_0$, the unit-level treatment effect:

$$F_{Y_1 - Y_0}(z) = \int \mathbb{1}(y_1 - y_0 \leq z) dF_{Y_1, Y_0}(y_1, y_0).$$

This parameter depends on the structure of the dependence between the two potential outcomes, which is unknown from either version of the unconfoundedness assumption. As discussed in Fan and Park (2010), $F_{Y_1 - Y_0}(z)$ is partially identified and sharp bounds can be recovered in terms of the distribution of (Y, X, W) .

To help classify treatment effect parameters, consider the decomposition

$$F_{Y_1, Y_0|X, W}(y_1, y_0 \mid x, w) = C_{1,0|X, W} \left(F_{Y_1|X, W}(y_1 \mid x, w), F_{Y_0|X, W}(y_0 \mid x, w) \mid x, w \right),$$

where $C_{1,0|X, W}(\cdot, \cdot \mid x, w)$ is a copula that characterizes the dependence between Y_1 and Y_0 conditional on $(X, W) = (x, w)$. By Sklar's Theorem (Sklar 1959), such a copula exists. Given that $F_{Y, X, W}$ is known, we consider treatment effect parameters that can be written as a function of $(F_{Y_1|X, W}, F_{Y_0|X, W}, C_{1,0|X, W}, F_{Y, X, W})$. We denote these parameters through the

functional

$$\theta(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}), \quad (2)$$

and we give several examples below. The dependence of θ on some of its arguments is suppressed if the functional is constant with respect to them. We will sometimes denote these parameters as functionals of $(F_{Y_1|W}, F_{Y_0|W}, F_{Y,X,W})$ rather than $(F_{Y_1|X,W}, F_{Y_0|X,W}, F_{Y,X,W})$. These two formulations are equivalent due to the relationship

$$F_{Y_x|W}(y | w) = F_{Y|X,W}(y | x, w)p_{x|w} + F_{Y_x|X,W}(y | 1 - x, w)p_{1-x|w} \quad (3)$$

which holds for all $(y, x, w) \in \mathbb{R} \times \{0, 1\} \times \text{supp}(W)$.

Next we consider eleven example target parameters. Our results give sharp bounds for all eleven parameters under our general class of assumptions, including the marginal sensitivity model as a special case. For many of these parameters we also obtain analytical, closed form expressions for the bound functions.

Example 1 (Conditional Quantile Treatment Effect). For quantile index $\tau \in (0, 1)$ and covariate value $w \in \text{supp}(W)$, the conditional quantile treatment effect (CQTE) can be written as

$$\text{CQTE}(\tau | w) := Q_{Y_1|W}(\tau | w) - Q_{Y_0|W}(\tau | w) = \theta_{\text{CQTE}}(F_{Y_1|W}, F_{Y_0|W}; \tau, w)$$

where $\theta_{\text{CQTE}}(F_{Y_1|W}, F_{Y_0|W}; \tau, w) := F_{Y_1|W}^{-1}(\tau | w) - F_{Y_0|W}^{-1}(\tau | w)$ where $F_{Y_x|W}^{-1}(\tau | w) = \inf\{y \in \mathbb{R} : F_{Y_x|W}(y | w) \geq \tau\}$ is the left-inverse of $F_{Y_x|W}(\cdot | w)$, or its conditional quantile function.

Example 2 (Conditional Average Treatment Effect). For covariate value $w \in \text{supp}(W)$, the conditional average treatment effect (CATE) can be written as

$$\text{CATE}(w) := \mathbb{E}[Y_1 - Y_0 | W = w] = \theta_{\text{CATE}}(F_{Y_1|W}, F_{Y_0|W}; w)$$

where $\theta_{\text{CATE}}(F_{Y_1|W}, F_{Y_0|W}; w) := \int y_1 dF_{Y_1|W}(y_1 | w) - \int y_0 dF_{Y_0|W}(y_0 | w)$.

Example 3 (Average Treatment Effect). Denote the average treatment effect (ATE) as

$$\text{ATE} := \mathbb{E}[Y_1 - Y_0] = \theta_{\text{ATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W)$$

where $\theta_{\text{ATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W) := \int (\int y_1 dF_{Y_1|W}(y_1 | w) - \int y_0 dF_{Y_0|W}(y_0 | w)) dF_W(w)$. We

can also consider weighted average treatment effects of the kind

$$\text{WATE}(\omega) := \mathbb{E}[\omega(W)(Y_1 - Y_0)]$$

for an identified function $\omega : \text{supp}(W) \rightarrow \mathbb{R}_{\geq 0}$. The ATE is a special case where $\omega(w) = 1$.

Example 4 (Average Treatment Effect on the Treated). Denote the average treatment effect on the treated (ATT) as

$$\text{ATT} := \mathbb{E}[Y_1 - Y_0 \mid X = 1] = \theta_{\text{ATT}}(F_{Y_0|X,W}, F_{Y,X,W})$$

where $\theta_{\text{ATT}}(F_{Y_0|X,W}, F_{Y,X,W}) := \mathbb{E}[Y \mid X = 1] - \int \int y_0 dF_{Y_0|X,W}(y_0 \mid 1, w) dF_{W|X}(w \mid 1)$.

Example 5 (Quantile Treatment Effect). For $\tau \in (0, 1)$, denote the quantile treatment effect (QTE) as

$$\text{QTE}(\tau) := Q_{Y_1}(\tau) - Q_{Y_0}(\tau) = \theta_{\text{QTE}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \tau)$$

where $\theta_{\text{QTE}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \tau) := F_{Y_1}^{-1}(\tau) - F_{Y_0}^{-1}(\tau)$ and $F_{Y_x}(\cdot) = \int F_{Y_x|W}(\cdot \mid w) dF_W(w)$ for $x \in \{0, 1\}$.

Example 6 (Quantile Treatment Effect on the Treated). For $\tau \in (0, 1)$, denote the quantile treatment effect on the treated (QTT) as

$$\text{QTT}(\tau) := Q_{Y_1|X}(\tau \mid 1) - Q_{Y_0|X}(\tau \mid 1) = \theta_{\text{QTT}}(F_{Y_0|X,W}, F_{Y,X,W}; \tau)$$

where $\theta_{\text{QTT}}(F_{Y_0|X,W}, F_{Y,X,W}; \tau) := Q_{Y|X}(\tau \mid 1) - F_{Y_0|X}^{-1}(\tau \mid 1)$ with

$$F_{Y_0|X}(\cdot \mid 1) = \int F_{Y_0|X,W}(\cdot \mid 1, w) dF_{W|X}(w \mid 1).$$

Example 7 (CATE Distribution). For $\tau \in (0, 1)$, denote the quantile of the CATE (QCATE) as

$$\text{QCATE}(\tau) := F_{\text{CATE}(W)}^{-1}(\tau) = \theta_{\text{QCATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \tau)$$

where $\theta_{\text{QCATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \tau) := F_{\text{CATE}(W)}^{-1}(\tau)$ with

$$F_{\text{CATE}(W)}(z) = \int \mathbb{1}(\theta_{\text{CATE}}(F_{Y_1|W}, F_{Y_0|W}; w) \leq z) dF_W(w).$$

This parameter is motivated by the sorted effects studied in Chernozhukov, Fernández-Val, and Luo (2018).

Example 8 (Average Weighted Welfare). For a weight (or assignment) function $\omega : \text{supp}(W) \rightarrow [0, 1]$, denote the average weighted welfare (AWW) as

$$\text{AWW}(\omega) := \mathbb{E}[\omega(W)Y_1 + (1 - \omega(W))Y_0] = \theta_{\text{AWW}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \omega)$$

where

$$\theta_{\text{AWW}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \omega) := \int \left(\omega(w) \int y_1 dF_{Y_1|W}(y_1 | w) + (1 - \omega(w)) \int y_0 dF_{Y_0|W}(y_0 | w) \right) dF_W(w).$$

Kallus and Zhou (2018) called the AWW the *policy value*.

Example 9 (Joint Distribution Function). For $(y_1, y_0) \in \mathbb{R}^2$, the joint cumulative distribution function (cdf) of (Y_1, Y_0) is

$$F_{Y_1, Y_0}(y_1, y_0) := \mathbb{P}(Y_1 \leq y_1, Y_0 \leq y_0) = \theta_{\text{CDF}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; y_1, y_0)$$

where

$$\begin{aligned} \theta_{\text{CDF}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; y_1, y_0) := \\ \sum_{x \in \{0,1\}} \int C_{1,0|X,W}(F_{Y_1|X,W}(y_1 | x, w), F_{Y_0|X,W}(y_0 | x, w) | x, w) p_{x|w} dF_W(w). \end{aligned}$$

Example 10 (Distribution of Treatment Effects). For $z \in \mathbb{R}$, the cumulative distribution function for the unit level treatment effect $Y_1 - Y_0$ (called the DTE) is

$$\text{DTE}(z) := F_{Y_1 - Y_0}(z) = \mathbb{P}(Y_1 - Y_0 \leq z) = \theta_{\text{DTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; z)$$

where

$$\begin{aligned} \theta_{\text{DTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; z) := \\ \sum_{x \in \{0,1\}} \int \left(\int_{\{y_1 - y_0 \leq z\}} dC_{1,0|X,W}(F_{Y_1|X,W}(y_1 | x, w), F_{Y_0|X,W}(y_0 | x, w) | x, w) \right) p_{x|w} dF_W(w). \end{aligned}$$

Example 11 (Quantiles of Treatment Effects). For $\tau \in (0, 1)$, the quantiles of the distribu-

tion function of treatment effect (QDTE) $Y_1 - Y_0$ is

$$Q_{Y_1 - Y_0}(\tau) := \inf\{z \in \mathbb{R} : F_{Y_1 - Y_0}(z) \geq \tau\} = \theta_{\text{QDTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; \tau)$$

where

$$\begin{aligned} &\theta_{\text{QDTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; \tau) \\ &:= \inf\{z \in \mathbb{R} : \theta_{\text{DTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; z) \geq \tau\}. \end{aligned}$$

The parameters in examples 1–8 only depend on the distribution of potential outcomes through their marginal distributions given (X, W) , while the parameters in examples 9–11 also depend on their copulas. Under overlap and unconfoundedness, the parameters in 1–8 are all point-identified. The parameters 9–11 are partially identified under overlap and unconfoundedness since the conditional copulas $C_{1,0|X,W}$ are not identified from the joint distribution of (Y, X, W) . In other words, these parameters depend on the type of dependence between Y_1 and Y_0 , and unconfoundedness does not reveal any information about this dependence. For example, Fan and Park (2010) show the identified set for $F_{Y_1 - Y_0}(z)$, the DTE, is an interval and they provide a closed-form expression for its lower and upper bounds.

However, if unconfoundedness fails, all these parameters will be partially identified. The identified set for parameters that are partially identified under unconfoundedness becomes larger under failures of unconfoundedness.

3 Relaxing Unconfoundedness

We now consider relaxations of the unconfoundedness assumption. We will consider two related, general relaxations of unconfoundedness that encompass several disparate relaxations that were studied in the literature. We begin by considering a class of assumptions on the probabilities of treatment when conditioning on covariates W and one of the potential outcomes.

Assumption 2 (Marginal c -dependence). Let $(\underline{c}(w, \eta), \bar{c}(w, \eta))$ satisfy $0 < \underline{c}(w, \eta) \leq p_{1|w} \leq \bar{c}(w, \eta) < 1$ for all $w \in \text{supp}(W)$. The potential outcomes satisfy *marginal c -dependence* if, for $x \in \{0, 1\}$,

$$p_x(Y_x, w) := \mathbb{P}(X = 1 \mid Y_x, W = w) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$.

This assumption restricts the manner in which potential outcomes affect the treatment probability $p_x(y, w)$, which we call a *latent propensity score*. We will use c -dependence assumptions to conduct sensitivity analyses for unconfoundedness. Here the sensitivity parameters are $\underline{c}(w, \eta)$ and $\bar{c}(w, \eta)$, which we refer to as bound functions. Like the notation in Rambachan et al. (2023), we let η be a possibly infinite-dimensional nuisance parameter that is point-identified from the distribution $F_{Y,X,W}$. The bound functions are also allowed to depend on the covariate value w . In principle, we can also allow the bounds to differ across $x \in \{0, 1\}$, but we do not include an x subscript for simplicity. The specification of $\underline{c}(w, \eta)$ and $\bar{c}(w, \eta)$ is left implicit, which allows them to be functions of low-dimensional or scalar sensitivity parameters. For example, the marginal sensitivity model of Tan (2006), which depends on a single sensitivity parameter, can be viewed as a special case of marginal c -dependence. We show this in section 3.1.

We can also see that setting $\underline{c}(w, \eta) = \bar{c}(w, \eta) = p_{1|w}$ yields unconfoundedness as a special case of marginal c -dependence, while letting $(\underline{c}(w, \eta), \bar{c}(w, \eta))$ approach $(0, 1)$ implies that no restrictions on the dependence between X and Y_x (given covariates) are imposed. Note that we restrict the propensity score $p_{1|w}$ to lie within our specified bounds for $p_x(Y_x, w)$. If the propensity score were outside these bounds, then the assumption would be misspecified because, by the law of iterated expectations, $p_{1|w} = \mathbb{E}[\mathbb{P}(X = 1 \mid Y_x, W = w) \mid W = w] \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$.

We also consider a closely related class of assumptions that restricts the probability of treatment given both potential outcomes.

Assumption 3 (Joint c -dependence). Let $(\underline{c}(w, \eta), \bar{c}(w, \eta))$ satisfy $0 < \underline{c}(w, \eta) \leq p_{1|w} \leq \bar{c}(w, \eta) < 1$ for all $w \in \text{supp}(W)$. The potential outcomes satisfy *joint c -dependence* if

$$p(Y_1, Y_0, w) := \mathbb{P}(X = 1 \mid Y_1, Y_0, W = w) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$.

Joint c -dependence with bound functions $\underline{c}(w, \eta)$ and $\bar{c}(w, \eta)$ implies marginal c -dependence with the same bound functions. This is due to the law of iterated expectations.

Lemma 1. Let Assumption 3 hold for $(\underline{c}(w, \eta), \bar{c}(w, \eta))$. Then, Assumption 2 holds for $(\underline{c}(w, \eta), \bar{c}(w, \eta))$.

We next show that several unconfoundedness relaxations from recent related literature can be viewed as special cases of either marginal or joint c -dependence.

3.1 The Marginal Sensitivity Model

Tan (2006) proposed the Marginal Sensitivity Model (MSM), which restricts the odds ratio between propensity scores and treatment probabilities that also condition on the potential outcome Y_x , for $x = 0, 1$. For $x \in \{0, 1\}$, let

$$R_x(Y_x, W) := \frac{\mathbb{P}(X = 1 \mid Y_x, W)}{\mathbb{P}(X = 0 \mid Y_x, W)} \bigg/ \frac{\mathbb{P}(X = 1 \mid W)}{\mathbb{P}(X = 0 \mid W)}$$

denote this odds ratio. When Y_x is continuously distributed with respect to the Lebesgue measure, this ratio can also be expressed as ratios of conditional densities of $Y_x \mid X = 1, W$ and $Y_x \mid X = 0, W$.

Tan (2006)'s MSM posits known bounds for these odds ratios.

Definition 1 (Marginal Sensitivity Model (MSM)). Let $\Lambda \in [1, +\infty)$ be known. The potential outcomes satisfy the *Marginal Sensitivity Model* if

$$R_x(Y_x, w) \in [\Lambda^{-1}, \Lambda] \text{ for } x \in \{0, 1\}$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$.

Λ is a scalar sensitivity parameter. Setting $\Lambda = 1$ is equivalent to assuming unconfoundedness, and increasing Λ allows for more dependence of latent propensity scores on potential outcomes. Variants of Tan (2006)'s MSM whose odds ratios condition on both potential outcomes have also been considered. Similarly, these odds ratios may instead condition on an abstract confounder U rather than potential outcomes. See, for example, Dorn and Guo (2023) and Dorn et al. (2024) for recent examples of these two variants. To distinguish it from the case where one conditions on the potential outcomes one at a time, we call the version that conditions on both potential outcomes the *Joint Sensitivity Model* (JSM).

Definition 2 (Joint Sensitivity Model (JSM)). Let $\Lambda \in [1, +\infty)$ be known. The potential outcomes satisfy the *Joint Sensitivity Model* if

$$R(Y_1, Y_0, w) \in [\Lambda^{-1}, \Lambda] \text{ for } x \in \{0, 1\}$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$, where

$$R(Y_1, Y_0, W) := \frac{\mathbb{P}(X = 1 \mid Y_1, Y_0, W)}{\mathbb{P}(X = 0 \mid Y_1, Y_0, W)} \bigg/ \frac{\mathbb{P}(X = 1 \mid W)}{\mathbb{P}(X = 0 \mid W)}.$$

We now state generalizations of the MSM and JSM that allow their odds ratios to have

arbitrary bounds, as opposed to bounds that have product equal to 1. We also allow their bounds to depend on covariates or nuisance parameters. We will continue distinguishing between *marginal* sensitivity models, which condition on one potential outcome at a time, and *joint* sensitivity models, which condition on both potential outcomes.

Definition 3 (Generalized Sensitivity Models). Let $\underline{\Lambda}(w, \eta) \in (0, 1]$ and $\bar{\Lambda}(w, \eta) \in [1, +\infty)$ for all $w \in \text{supp}(W)$ where $\underline{\Lambda}(w, \eta)$ and $\bar{\Lambda}(w, \eta)$ are known. The potential outcomes satisfy the *Generalized Marginal Sensitivity Model* (GMSM) if

$$R_x(Y_x, w) \in [\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)] \text{ for } x \in \{0, 1\}$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$. They satisfy the *Generalized Joint Sensitivity Model* (GJSM) if

$$R(Y_1, Y_0, w) \in [\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)]$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$.

We can see that the MSM is a special case of the GMSM by setting $[\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)] = [\Lambda^{-1}, \Lambda]$. Similarly, the JSM is a special case of the GJSM.

The GMSM is equivalent to marginal c -dependence because, for each bound function pair $[\underline{c}(w, \eta), \bar{c}(w, \eta)]$ under marginal c -dependence, there exists exactly one corresponding bound function pair $[\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)]$ under the GMSM. The same link exists between joint c -dependence and the GJSM. We show this in the following proposition.

Proposition 1 (Equivalence of Sensitivity Models).

1. Let marginal (joint) c -dependence hold with bound functions $[\underline{c}(w, \eta), \bar{c}(w, \eta)]$. Then the GMSM (GJSM) holds with bound functions

$$[\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)] = \left[\frac{\underline{c}(w, \eta) p_{0|w}}{1 - \underline{c}(w, \eta) p_{1|w}}, \frac{\bar{c}(w, \eta) p_{0|w}}{1 - \bar{c}(w, \eta) p_{1|w}} \right]. \quad (4)$$

2. Let the GMSM (GJSM) hold with bound functions $[\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)]$. Then marginal (joint) c -dependence holds with bound functions

$$[\underline{c}(w, \eta), \bar{c}(w, \eta)] = \left[\frac{p_{1|w} \underline{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w} \underline{\Lambda}(w, \eta)}, \frac{p_{1|w} \bar{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w} \bar{\Lambda}(w, \eta)} \right]. \quad (5)$$

This proposition shows that the marginal c -dependence is equivalent to the generalized marginal sensitivity model. Similarly, joint c -dependence is equivalent to the generalized marginal sensitivity model.

3.2 Conditional c -dependence

Masten and Poirier (2018a) studied a relaxation of unconfoundedness they called *conditional c -dependence*, which assumed symmetric bounds on the latent propensity score.

Definition 4 (Conditional c -dependence). Let $c \in [0, 1]$ be a known scalar sensitivity parameter. The potential outcomes satisfy *conditional c -dependence* if

$$p_x(Y_x, w) := \mathbb{P}(X = 1 \mid Y_x, W = w) \in [p_{1|w} - c, p_{1|w} + c]$$

almost surely conditional on $W = w$ for all $w \in \text{supp}(W)$.

This is a special case of marginal c -dependence where the bounds equal

$$\underline{c}(w, \eta) = p_{1|w} - c \quad \text{and} \quad \bar{c}(w, \eta) = p_{1|w} + c.$$

Here the nuisance parameter is $p_{1|(\cdot)}$, the propensity score function. Unconfoundedness is obtained by setting $c = 0$, while the no-assumption bounds are obtained for c 's equal to or larger than $\sup_{w \in \text{supp}(W)} \max\{p_{1|w}, p_{0|w}\}$. Masten and Poirier (2018a) provided closed-form expressions for sharp bounds on the CQTE, CATE, ATE, QTE, and ATT when potential outcomes are continuously distributed or binary. Masten, Poirier, and Zhang (2024) describe flexible parametric estimators of these bounds and provide nonstandard inference methods.

4 General Identification Results

Next we derive sharp bounds on a large class of target parameters under the relaxations described in section 3. We will study a class of parameters that includes all eleven examples in section 2 as special cases. Specifically, we will compute these parameters' sharp bounds, or their identified set, under marginal and joint c -dependence, which are equivalent to the GMSM and GJSM, respectively.

4.1 Bounds on Marginal Distributions

Before studying our general class of treatment effect parameters, we first consider bounds on the distribution functions of each potential outcome, given covariates. These cdfs are

building blocks for these parameters and, as we will see, analytical bounds on these cdfs will directly map into analytical bounds on these parameters.

Specifically, we begin by analyzing the conditional cdf of the potential outcome Y_x given the covariate value w , $F_{Y_x|W}(y | w) := \mathbb{P}(Y_x \leq y | W = w)$. Under either marginal or joint c -dependence, we can show that this cdf is bounded above and below by two cdfs which form an envelope for $F_{Y_x|W}(y | w)$ for all values of $(y, w) \in \mathbb{R} \times \text{supp}(W)$.

Define the following functions:

$$\begin{aligned} \underline{F}_{Y_1|W}(y | w) &= \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)}, \frac{\underline{c}(w, \eta) - p_{1|w}}{\underline{c}(w, \eta)} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)} \right\} \\ \overline{F}_{Y_1|W}(y | w) &= \min \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)}, \frac{\overline{c}(w, \eta) - p_{1|w}}{\overline{c}(w, \eta)} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\overline{c}(w, \eta)} \right\} \\ \underline{F}_{Y_0|W}(y | w) &= \max \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \underline{c}(w, \eta)}, \frac{p_{1|w} - \overline{c}(w, \eta)}{1 - \overline{c}(w, \eta)} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \overline{c}(w, \eta)} \right\} \\ \overline{F}_{Y_0|W}(y | w) &= \min \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \overline{c}(w, \eta)}, \frac{p_{1|w} - \underline{c}(w, \eta)}{1 - \underline{c}(w, \eta)} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \underline{c}(w, \eta)} \right\}. \end{aligned}$$

Viewed as functions of y , these four functions are cdfs since they are nondecreasing, right-continuous, and their limits as $y \rightarrow -\infty, +\infty$ equal 0 and 1, respectively. We show these four cdfs form bounds for $F_{Y_x|W}$ under marginal or joint c -dependence.

Lemma 2. Let Assumption 1 hold. Let either Assumption 2 or 3 hold. Then, for all $(y, w) \in \mathbb{R} \times \text{supp}(W)$,

$$\mathbb{P}(Y_1 \leq y | W = w) \in [\underline{F}_{Y_1|W}(y | w), \overline{F}_{Y_1|W}(y | w)]$$

and

$$\mathbb{P}(Y_0 \leq y | W = w) \in [\underline{F}_{Y_0|W}(y | w), \overline{F}_{Y_0|W}(y | w)].$$

We note a few properties of these bounds. The bounds for $F_{Y_x|W}(y | w)$ collapse to a point if either $\underline{c}(w, \eta) = p_{1|w}$ or $\overline{c}(w, \eta) = p_{1|w}$. Also note that $F_{Y|X,W}(y | x, w)$ always lies within the bounds for $F_{Y_x|W}(y | w)$. This is because c -dependence never rules out unconfoundedness, and unconfoundedness implies that the distribution of Y given $(X, W) = (x, w)$ equals that of Y_x given $W = w$.

These cdf bounds also yield cdf bounds under the GSM or GJM since they are equivalent to c -dependence as shown in Proposition 1. These bounds are also valid for the standard MSM or JM, as they are special cases of marginal or joint c -dependence.

We now show these cdf bounds are sharp, or that they cannot be improved upon. This is the case under marginal or joint c -dependence. The cdf of $Y_x | W = w$ can also lie in the

interior of these bounds, as we show that any convex linear combination of the upper and lower cdf bounds can be attained.

Before establishing this, let \mathcal{C} denote the set of all bivariate copulas and let

$$\mathcal{C}_{1,0|X,W} = \left\{ \{C_{1,0|x,w}\}_{x \in \{0,1\}, w \in \text{supp}(W)} \text{ such that } C_{1,0|x,w} \in \mathcal{C} \right\}$$

denote the collection of all bivariate copulas across all treatment and covariate values $(x, w) \in \{0, 1\} \times \text{supp}(W)$. We also say that a distribution function for $(Y_1, Y_0) | X, W$ is compatible with the observed distribution $F_{Y,X,W}$ if

$$F_{Y_1|X,W}(\cdot | 1, w) = F_{Y|X,W}(\cdot | 1, w) \quad \text{and} \quad F_{Y_0|X,W}(\cdot | 0, w) = F_{Y|X,W}(\cdot | 0, w) \quad (6)$$

for all $w \in \text{supp}(W)$.

We now define the identified set for the distribution of $(Y_1, Y_0) | X, W$ from the observable distribution $F_{Y,X,W}$ under c -dependence. This set consists of all conditional cdfs and copulas that imply a distribution for $(Y_1, Y_0) | X, W$ that is both compatible with the data distribution $F_{Y,X,W}$ and with a c -dependence condition.

Definition 5 (Identified Set). For a given distribution of the observables $F_{Y,X,W}$ and bound functions $c := (\underline{c}(w, \eta), \bar{c}(w, \eta))$, the identified set for $(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W})$ under marginal c -dependence is given by

$$\begin{aligned} \mathcal{I}^{\text{marg}}(F_{Y,X,W}; c) := & \{(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}) : F_{Y_1, Y_0|X,W} = C_{1,0|X,W}(F_{Y_1|X,W}, F_{Y_0|X,W}) \\ & \text{and } p_{1|(\cdot)} \text{ satisfy equation (6) and Assumption 2}\}. \end{aligned}$$

The identified set under joint c -dependence is given by

$$\begin{aligned} \mathcal{I}^{\text{joint}}(F_{Y,X,W}; c) := & \{(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}) : F_{Y_1, Y_0|X,W} = C_{1,0|X,W}(F_{Y_1|X,W}, F_{Y_0|X,W}) \\ & \text{and } p_{1|(\cdot)} \text{ satisfy equation (6) and Assumption 3}\}. \end{aligned}$$

In our later derivations, we sometimes refer to the identified set for $(F_{Y_1|W}, F_{Y_0|W}, C_{1,0|X,W})$ instead, which we denote by $\mathcal{I}_0^i(F_{Y,X,W}; c)$ for $i \in \{\text{marg}, \text{joint}\}$. Via equation (3), this set can be viewed as an affine transformation of the identified set for $(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W})$.

We now show some key properties of the cdfs and copulas in these identified sets. We begin with marginal c -dependence.

Theorem 1. Let Assumptions 1 and 2 hold. For all $(\varepsilon, \gamma) \in [0, 1]^2$ and for any $C_{1,0|X,W} \in$

$\mathcal{C}_{1,0|X,W}$,

$$(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \overline{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \overline{F}_{Y_0|W}, C_{1,0|X,W}) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c).$$

This theorem shows that the four pairs of cdfs $(\underline{F}_{Y_1|W}, \underline{F}_{Y_0|W})$, $(\overline{F}_{Y_1|W}, \underline{F}_{Y_0|W})$, $(\underline{F}_{Y_1|W}, \overline{F}_{Y_0|W})$, and $(\overline{F}_{Y_1|W}, \overline{F}_{Y_0|W})$ are part of the identified set. This is obtained by varying (ε, γ) over $\{(1, 1), (0, 1), (1, 0), (0, 0)\}$. We show this by explicitly constructing latent propensity scores $p_1(Y_1, w)$ and $p_0(Y_0, w)$ that lie in $[\underline{c}(w, \eta), \overline{c}(w, \eta)]$ almost surely under the implied distribution of $(Y_1, Y_0) | W = w$. These propensity scores have a switching structure where they equal the lower/upper bound for low values of Y_x and the upper/lower bound for large values of Y_x . For example, the propensity score $p_1(Y_1, w)$ associated with cdf upper bound $\overline{F}_{Y_1|W}$ equals

$$p_1(Y_1, w) := \begin{cases} \underline{c}(w, \eta) & \text{if } Y_1 < \overline{Q}_1 \\ \overline{A}_1 & \text{if } Y_1 = \overline{Q}_1 \\ \overline{c}(w, \eta) & \text{if } Y_1 > \overline{Q}_1 \end{cases}$$

where

$$\overline{Q}_1 := Q_{Y|X,W} \left(\frac{(\overline{c}(w, \eta) - p_{1|w}) \underline{c}(w, \eta)}{(\overline{c}(w, \eta) - \underline{c}(w, \eta)) p_{1|w}} \mid X = 1, W = w \right)$$

and

$$\overline{A}_1 := \frac{\mathbb{P}(Y = \overline{Q}_1, X = 1 \mid W = w)}{\overline{F}_{Y_1|W}(\overline{Q}_1 \mid w) - \overline{F}_{Y_1|W}(\overline{Q}_1 - \mid w)}.$$

Note that $\overline{A}_1 \in [\underline{c}(w, \eta), \overline{c}(w, \eta)]$. We denote $\lim_{q \nearrow \overline{Q}_1} \overline{F}_{Y_1|W}(q \mid w)$ by $\overline{F}_{Y_1|W}(\overline{Q}_1 - \mid w)$. The propensity scores associated with the cdf bounds $\underline{F}_{Y_1|W}$, $\overline{F}_{Y_0|W}$, and $\underline{F}_{Y_0|W}$ can all be found in Appendix A.

This switching structure of the latent propensity score was observed for conditional c -dependence by Masten and Poirier (2018a, pages 335–339), and for the MSM in Proposition 2 of Dorn and Guo (2023). Our sharpness proof implies that latent propensity score $p_1(Y_1, w)$ satisfies an integral constraint, namely that $\mathbb{E}[p_1(Y_1, W) \mid W = w] = \mathbb{P}(X = 1 \mid W = w)$, in order to ensure it is compatible with the observed propensity score.

Theorem 1 also shows that any convex linear combinations of these four cdf pairs lies in the identified set. As a result, the identified set for $F_{Y_x|W}(y \mid w)$ is the entire closed interval $[\underline{F}_{Y_x|W}(y \mid w), \overline{F}_{Y_x|W}(y \mid w)]$. Moreover, the identified set for the pair $(F_{Y_1|W}(y \mid w),$

w), $F_{Y_0|W}(y | w)$) is the Cartesian product of their individual identified sets, meaning that fixing or knowing the conditional distribution of one potential outcome does not affect the identified set of the distribution of the other potential outcome.

Finally, Theorem 1 proves that no conditional copulas are ruled out by marginal c -dependence. For example, marginal c -dependence allows Y_1 and Y_0 to be independent, comonotonic¹, or countermonotonic given X and W .

A similar result is obtained under joint c -dependence.

Theorem 2. Let Assumptions 1 and 3 hold. For all $(\varepsilon, \gamma) \in [0, 1]^2$ there exists $C_{1,0|X,W} \in \mathcal{C}_{1,0|X,W}$ such that

$$(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \overline{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \overline{F}_{Y_0|W}, C_{1,0|X,W}) \in \mathcal{I}_0^{\text{joint}}(F_{Y,X,W}; c).$$

The only difference between the two theorems concerns the dependence structures between Y_1 and Y_0 . Theorem 1 shows that all copulas are compatible with marginal c -dependence, while our proof of Theorem 2 only exhibits one copula for each pair of conditional distributions $(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \overline{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \overline{F}_{Y_0|W})$.

4.2 Bounds on Monotonic Parameters

The sharp bounds provided in theorems 1 and 2 can be used to deliver analytical expressions for sharp bounds on a large class of treatment effect parameters. We first define the identified set for a parameter θ defined in (2).

Definition 6 (Parameter Identified Sets). Let $\theta(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W})$ be a parameter. Its identified set under marginal c -dependence with bounds $c := (\underline{c}(w, \eta), \overline{c}(w, \eta))$ is given by

$$\mathcal{I}_\theta^{\text{marg}}(F_{Y,X,W}; c) := \{\theta(F_1, F_0, C, F_{Y,X,W}) : (F_1, F_0, C) \in \mathcal{I}^{\text{marg}}(F_{Y,X,W}; c)\}.$$

For a parameter $\theta(F_{Y_1|X,W}, F_{Y_0|X,W}, F_{Y,X,W})$ that does not depend on the copula $C_{1,0|X,W}$, its identified set under joint c -dependence is given by

$$\mathcal{I}_\theta^j(F_{Y,X,W}; c) := \{\theta(F_1, F_0, F_{Y,X,W}) : (F_1, F_0, C) \in \mathcal{I}^{\text{joint}}(F_{Y,X,W}; c) \text{ for some } C \in \mathcal{C}_{1,0|X,W}\}.$$

These sets are the parameter values consistent with the known distribution of observables $F_{Y,X,W}$ and with a c -dependence condition. Without restrictions on how θ depends on the

¹This is also referred to as *rank invariance*. For example, see the discussion in Heckman, Smith, and Clements (1997).

distribution of potential outcomes, these sets may take various shapes.

We focus on a class of scalar estimands that can be ordered with respect to first order stochastic dominance.

Definition 7 (First-Order Stochastic Dominance). Let \mathcal{F} be the set of all univariate cdfs and let $F, G \in \mathcal{F}$. Say that F first-order stochastically dominates G , denoted by $F \succeq G$, if $F(u) \leq G(u)$ for all $u \in \mathbb{R}$.

Next we define our target class of parameters.

Definition 8 (Monotonic Parameters). Let $\theta : \mathcal{F} \rightarrow \mathbb{R}$ be a parameter. Say that θ is *increasing* if $F \succeq G$ implies $\theta(F) \geq \theta(G)$. Say that θ is *decreasing* if $-\theta$ is increasing, and say θ is *monotonic* if it is either increasing or decreasing.

Following Manski (1997), monotonic parameters are also called D -parameters, or D_1 -parameters. Also see Manski (2003) or Stoye (2010) who consider parameters that are increasing with respect to second-order stochastic dominance.

As an example, consider a parameter $\theta(F_{Y_1|W})$ that is increasing in $F_{Y_1|W}$ and suppose c -dependence holds. Then

$$\theta(F_{Y_1|W}) \in [\theta(\overline{F}_{Y_1|W}), \theta(\underline{F}_{Y_1|W})]$$

since $\overline{F}_{Y_1|W} \preceq F_{Y_1|W} \preceq \underline{F}_{Y_1|W}$, which holds by Lemma 2. This interval cannot be made narrower since the cdf bounds $[\underline{F}_{Y_1|W}, \overline{F}_{Y_1|W}]$ are sharp by theorems 1 and 2. Therefore, the identified set for $\theta(F_{Y_1|W})$ is a subset of this closed interval that always contains its two endpoints. The interior of this interval is also part of the identified set if the functional θ is continuous in the sense that $\varepsilon \mapsto \theta(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \overline{F}_{Y_1|W})$ is continuous. This type of continuity is implied by the continuity of the mapping $F \mapsto \theta(F)$ under the sup-distance metric.

Assuming monotonicity of the parameter will help derive properties of its identified set. Monotonicity is a substantive restriction, but all eleven parameters from Section 2 satisfy it. This is formally established in Lemma 3 below. We begin by considering monotonic parameters that do not depend on copulas.

Theorem 3. Let $\theta(F_{Y_1|W}, F_{Y_0|W}, F_{Y,X,W})$ be increasing in $F_{Y_1|W}(\cdot | w)$ and decreasing in $F_{Y_0|W}(\cdot | w)$ for each $w \in \text{supp}(W)$. Let Assumption 1 hold, and either Assumption 2 or 3 hold. Then, for $i \in \{\text{marg}, \text{joint}\}$ the convex hull of the identified set for $\theta(F_{Y_1|W}, F_{Y_0|W}, F_{Y,X,W})$

is the closed interval

$$\begin{aligned} \mathcal{I}_\theta^i(F_{Y,X,W}; c) &= \left[\inf_{(F_1, F_0, C) \in \mathcal{I}_0^i(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}), \sup_{(F_1, F_0, C) \in \mathcal{I}_0^i(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) \right] \\ &= [\theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W}), \theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W})]. \end{aligned}$$

If $(\varepsilon, \gamma) \mapsto \theta(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \bar{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \bar{F}_{Y_0|W}, F_{Y,X,W})$ is continuous over $(\varepsilon, \gamma) \in [0, 1]^2$, this interval equals the identified set.

This theorem shows that substituting the upper/lower cdf bounds delivers sharp bounds for any parameter that is monotonic in the first-order stochastic dominance sense. The result is derived under an assumption that the parameter is increasing in $F_{Y_1|W}$ and decreasing in $F_{Y_0|W}$, but it immediately generalizes to parameters that are increasing or decreasing in either or both conditional cdfs. For example, the cdf pair $(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W})$ will maximize a parameter that is decreasing in $F_{Y_1|W}$ and increasing in $F_{Y_0|W}$, and the cdf pair $(\bar{F}_{Y_1|W}, \bar{F}_{Y_0|W})$ will maximize (minimize) a parameter that is decreasing (increasing) in both $F_{Y_1|W}$ and $F_{Y_0|W}$. The identified set for these parameters always contains endpoints $\theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W})$ and $\theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W})$. It also contains all the values between these endpoints whenever the mapping θ is continuous in the appropriate sense.

We document the monotonicity of various building blocks for parameters of interest in the following technical lemma. We omit covariates W for simplicity here, except in part 4 on QCATE because that parameter requires covariates to be nontrivial.

Lemma 3. Let Assumption 1 hold. Then, for $x \in \{0, 1\}$ and $\tau \in (0, 1)$,

1. $\theta_{\mathbb{E}}(F_{Y_x}) := \int y dF_{Y_x}(y)$ is increasing and continuous in the sense that $\varepsilon \mapsto \theta_{\mathbb{E}}(\varepsilon F_{Y_x} + (1 - \varepsilon) F'_{Y_x})$ is continuous for any (F_{Y_x}, F'_{Y_x}) over $\varepsilon \in [0, 1]$.
2. $\theta_Q(F_{Y_x}; \tau) := F_{Y_x}^{-1}(\tau)$ is increasing.
3. $\theta_{CQ}(F_{Y_x}; \tau) := F_{Y_x|X}^{-1}(\tau | 1 - x)$ is increasing.
4. $\theta_{\text{QCATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W; \tau)$ (see Example 7) is decreasing in $F_{Y_1|W}$ and increasing in $F_{Y_0|W}$.
5. $\theta_{\text{CDF}}(F_{Y_1}, F_{Y_0}, C; y_1, y_0) := C(F_{Y_1}(y_0), F_{Y_0}(y_0))$ is decreasing in F_{Y_1} and F_{Y_0} for all $(y_1, y_0) \in \mathbb{R}^2$ and copulas C .
6. $\theta_{\text{DTE}}(F_{Y_1}, F_{Y_0}, C; z) := \int_{\{y_1 - y_0 \leq z\}} dC(F_{Y_1}(y_1), F_{Y_0}(y_0))$ is decreasing in F_{Y_1} and increasing in F_{Y_0} for all $z \in \mathbb{R}$ and copulas C .

Using this lemma, all eight parameters that are invariant to copulas are bounded by substituting the upper or lower cdf bounds from Theorem 2. This allows us to compute analytical bounds for these parameters.

5 Analytical Bounds on Treatment Effect Parameters

We explore these analytical bounds by focusing on five of our examples to illustrate these expressions. The first three parameters are independent from the copula, while the last two are copula-dependent.

5.1 Average Treatment Effects (Example 3)

From Lemma 3.1, we have that the ATE satisfies

$$\mathbb{E}[Y_1 - Y_0] = \theta_{\text{ATE}}(F_{Y_1|W}, F_{Y_0|W}, F_W) \in [\theta_{\text{ATE}}(\overline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_W), \theta_{\text{ATE}}(\underline{F}_{Y_1|W}, \overline{F}_{Y_0|W}, F_W)].$$

This interval equals the identified set by the monotonicity and continuity of the expectation functional which was established in Lemma 3.1. The lower and upper bounds can be obtained by calculating $\int y d\overline{F}_{Y_x|W}(y | w)$ and $\int y d\underline{F}_{Y_x|W}(y | w)$ for $x \in \{0, 1\}$. Via the quantile transformation, these bounds can also be written as integrals of $\underline{Q}_{Y_x|W}(u | w)$ and $\overline{Q}_{Y_x|W}(u | w)$ over $u \in (0, 1)$. Thus the ATE bounds can be written as integrals of quantiles. Via Lemma 14 in Appendix E, we show that these quantile integrals can be converted into conditional expectations of outcomes given that they exceed or fall short of a fixed conditional quantile. These are equivalent to Conditional Value at Risk (CVaR) measures that appear in Dorn et al. (2024). These bounds are stated explicitly in equations (52)–(55) in Appendix E in the general case. When $Y | X, W$ is continuously distributed, we obtain simpler expressions for these bounds that we give here:

$$\begin{aligned} \theta_{\text{ATE}}(\overline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_W) &= \mathbb{E} \left[\left(\mathbb{E}[Y | Y \leq \overline{Q}_1, X = 1, W] - \mathbb{E}[Y | Y \leq \underline{Q}_0, X = 0, W] \right) \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} \right] \\ &\quad + \mathbb{E} \left[\left(\mathbb{E}[Y | Y > \overline{Q}_1, X = 1, W] - \mathbb{E}[Y | Y > \underline{Q}_0, X = 0, W] \right) \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} \right] \end{aligned}$$

and

$$\begin{aligned} \theta_{\text{ATE}}(\underline{F}_{Y_1|W}, \overline{F}_{Y_0|W}, F_W) &= \mathbb{E} \left[\left(\mathbb{E}[Y | Y \leq \underline{Q}_1, X = 1, W] - \mathbb{E}[Y | Y \leq \overline{Q}_0, X = 0, W] \right) \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} \right] \\ &\quad + \mathbb{E} \left[\left(\mathbb{E}[Y | Y > \underline{Q}_1, X = 1, W] - \mathbb{E}[Y | Y > \overline{Q}_0, X = 0, W] \right) \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} \right]. \end{aligned}$$

Note that the dependence of (\underline{c}, \bar{c}) on (W, η) was suppressed for convenience.

5.2 Quantile Treatment Effects (Example 5)

We now consider bounds on the quantile treatment effect for a fixed quantile $\tau \in (0, 1)$. By Lemma 3.2, the functional θ_{QTE} is increasing in $F_{Y_1|W}$ and decreasing in $F_{Y_0|W}$. Therefore, by Theorem 3, $\text{QTE}(\tau)$ has the following sharp bounds:

$$\text{QTE}(\tau) \in \left[\underline{Q}_{Y_1}(\tau) - \bar{Q}_{Y_0}(\tau), \bar{Q}_{Y_1}(\tau) - \underline{Q}_{Y_0}(\tau) \right]$$

where \bar{Q}_{Y_x} is the left-inverse of cdf $\underline{F}_{Y_x}(\cdot) := \mathbb{E}[\underline{F}_{Y_x|W}(\cdot | W)]$ for $x \in \{0, 1\}$. Analogously, \underline{Q}_{Y_x} is the left-inverse of cdf $\bar{F}_{Y_x}(\cdot) := \mathbb{E}[\bar{F}_{Y_x|W}(\cdot | W)]$. Analytical expressions for the unconditional cdf bounds for the treated potential outcome are given by

$$\begin{aligned} \bar{F}_{Y_1}(y) &= \mathbb{E} \left[\min \left\{ F_{Y|X,W}(y | 1, W) \frac{p_{1|W}}{\underline{c}}, \frac{\bar{c} - p_{1|W}}{\bar{c}} + F_{Y|X,W}(y | 1, W) \frac{p_{1|W}}{\bar{c}} \right\} \right] \\ \underline{F}_{Y_1}(y) &= \mathbb{E} \left[\max \left\{ F_{Y|X,W}(y | 1, W) \frac{p_{1|W}}{\bar{c}}, \frac{\underline{c} - p_{1|W}}{\underline{c}} + F_{Y|X,W}(y | 1, W) \frac{p_{1|W}}{\underline{c}} \right\} \right] \end{aligned}$$

and similar expressions can be obtained for Y_0 . The left-inverses of the previous expressions yield bounds on quantiles of Y_1 and Y_0 , and which can be used to compute the QTE bounds.

5.3 Average Weighted Welfare (Example 8)

Consider a policy $\omega : \text{supp}(W) \rightarrow [0, 1]$ that treats units with covariate value w with probability $\omega(w)$. The average welfare in a population under such policy is given by

$$\text{AWW}(\omega) = \theta_{\text{AWW}}(F_{Y_1|W}, F_{Y_0|W}, F_W, \omega) = \mathbb{E}[\omega(W)\mathbb{E}[Y_1 | W]] + (1 - \mathbb{E}[\omega(W)])\mathbb{E}[Y_0 | W].$$

By adapting Lemma 3.1, this functional is increasing in $F_{Y_1|W}$, increasing in $F_{Y_0|W}$, and continuous in the sense defined in the lemma. Therefore, by Theorem 3, its identified set is the closed interval given by

$$\left[\theta_{\text{AWW}}(\bar{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_W, \omega), \theta_{\text{AWW}}(\underline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_W, \omega) \right].$$

An analytical expression for these bounds can be obtained by substituting in the expressions for the cdf bounds in the previous functionals. When $Y | X, W$ is continuously distributed,

the bounds are given by

$$\begin{aligned} & \theta_{\text{AWW}}(\bar{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_W, \omega) \\ &= \mathbb{E} \left[\omega(W) \left(\mathbb{E}[Y | Y \leq \bar{Q}_1, X = 1, W] \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} + \mathbb{E}[Y | Y > \bar{Q}_1, X = 1, W] \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} \right) \right] \\ &+ \mathbb{E} \left[(1 - \omega(W)) \left(\mathbb{E}[Y | Y \leq \bar{Q}_0, X = 0, W] \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} + \mathbb{E}[Y | Y > \bar{Q}_0, X = 0, W] \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} \right) \right] \end{aligned}$$

and

$$\begin{aligned} & \theta_{\text{AWW}}(\underline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_W, \omega) \\ &= \mathbb{E} \left[\omega(W) \left(\mathbb{E}[Y | Y \leq \underline{Q}_1, X = 1, W] \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} + \mathbb{E}[Y | Y > \underline{Q}_1, X = 1, W] \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} \right) \right] \\ &+ \mathbb{E} \left[(1 - \omega(W)) \left(\mathbb{E}[Y | Y \leq \underline{Q}_0, X = 0, W] \frac{\bar{c} - p_{1|W}}{\bar{c} - \underline{c}} + \mathbb{E}[Y | Y > \underline{Q}_0, X = 0, W] \frac{p_{1|W} - \underline{c}}{\bar{c} - \underline{c}} \right) \right]. \end{aligned}$$

5.4 Copula-Dependent Parameters

We now consider identification of the parameters in examples 9–11 which all depend on the copulas $C_{1,0|X,W}$. Even under unconfoundedness these parameters are not point-identified. Relaxing unconfoundedness will yield larger identified sets for these parameters when compared to the unconfoundedness baseline. We will focus on marginal c -dependence since it does not restrict the dependence structure between the potential outcomes.

The Joint Distribution Function

Consider identification of the joint cdf $F_{Y_1, Y_0}(y_1, y_0)$ under marginal c -dependence. Consider the functional

$$\begin{aligned} & \theta_{\text{CDF}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; y_1, y_0) \\ &:= \int (C_{1,0|X,W}(F_{Y_1|X,W}(y_1 | 1, w), F_{Y_0|X,W}(y_0 | 1, w) | 1, w) p_{1|w} \\ &\quad + C_{1,0|X,W}(F_{Y_1|X,W}(y_1 | 0, w), F_{Y_0|X,W}(y_0 | 0, w) | 0, w) p_{0|w}) dF_W(w). \end{aligned}$$

Fix the conditional copula function $C_{1,0|X,W}(\cdot, \cdot | \cdot, \cdot)$. Then by Lemma 3.4, this functional is decreasing in $F_{Y_0|X,W}(y_0 | 1, w)$ and $F_{Y_1|X,W}(y_1 | 0, w)$. Thus it is bounded below by

$$\theta_{\text{CDF}}(\underline{F}_{Y_1|X,W}, \underline{F}_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; y_1, y_0)$$

and above by

$$\theta_{\text{CDF}}(\overline{F}_{Y_1|X,W}, \overline{F}_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; y_1, y_0).$$

Moreover, by Theorem 3, these bounds are sharp.

Since $C_{1,0|X,W}$ is unknown, we then compute the maximum and minimum of these bounds over the set of copulas that are consistent with marginal c -dependence; this is simply the set of all copulas. The Fréchet-Hoeffding bounds show that all copulas C satisfy

$$C(u, v) \in [\max\{u + v - 1, 0\}, \min\{u, v\}] =: [\underline{C}(u, v), \overline{C}(u, v)]$$

for all $(u, v) \in [0, 1]^2$. The copula bounds \underline{C} and \overline{C} are themselves copulas. Combining these facts, we obtain the following analytical bounds on the joint cdf of potential outcomes.

Proposition 2 (Identified set for joint cdf). Let assumptions 1 and 2 hold. Then, for any $(y_1, y_0) \in \mathbb{R}^2$, the identified set for $F_{Y_1, Y_0}(y_1, y_0)$ is given by the closed interval

$$\mathcal{I}_{\theta_{\text{CDF}}}^{\text{marg}}(F_{Y,X,W}; c) = \left[\mathbb{E} \left(\max\{ \underline{F}_{Y_1|X,W}(y_1 | X, W) + \underline{F}_{Y_0|X,W}(y_0 | X, W) - 1, 0 \} \right), \right. \\ \left. \mathbb{E} \left(\min\{ \overline{F}_{Y_1|X,W}(y_1 | X, W), \overline{F}_{Y_0|X,W}(y_0 | X, W) \} \right) \right]. \quad (7)$$

The bounds in (7) are themselves cdfs, so these bounds can be attained simultaneously for all $(y_1, y_0) \in \mathbb{R}^2$. The bounds for F_{Y_1, Y_0} under unconfoundedness are obtained as a special case when $\underline{c} = p_{1|W} = \overline{c}$, which implies that $\overline{F}_{Y_x|X,W} = \underline{F}_{Y_x|X,W} = F_{Y|X,W}(\cdot | x, \cdot)$. Making this substitution in equation (7) yields these bounds under unconfoundedness.

The Distribution of Treatment Effects

Identification of this parameter under unconfoundedness was studied in Fan and Park (2010), by applying results first shown in Makarov (1982) and later studied in Williamson and Downs (1990). Masten and Poirier (2020) also studied this parameter under conditional c -dependence and under a range of assumptions on copulas for (Y_1, Y_0) . By Lemma 2.1 in Fan and Park (2010), the cdf of $Y_1 - Y_0$ given $(X, W) = (x, w)$ satisfies

$$F_{Y_1 - Y_0|X,W}(z | x, w) \in \left[\max \left\{ \sup_{y \in \mathbb{R}} (F_{Y_1|X,W}(y | x, w) - F_{Y_0|X,W}(y - z | x, w)), 0 \right\}, \right. \\ \left. 1 + \min \left\{ \inf_{y \in \mathbb{R}} (F_{Y_1|X,W}(y | x, w) - F_{Y_0|X,W}(y - z | x, w)), 0 \right\} \right]$$

and these bounds are sharp for any pair of cdfs $(F_{Y_1|X,W}, F_{Y_0|X,W})$. These bounds are decreasing in $F_{Y_1|X,W}$ and increasing in $F_{Y_0|X,W}$, therefore substituting the upper/lower cdf bounds

for $F_{Y_x|X,W}$ results in sharp bounds for $F_{Y_1-Y_0|X,W}$ under c -dependence. This was established in Lemma 3.6. Integrating these bounds over the marginal distribution of (X, W) yields sharp bounds for the unconditional cdf of $Y_1 - Y_0$. This result is summarized in the following proposition.

Proposition 3 (Identified set for DTE). Let assumptions 1 and 2 hold. For any $z \in \mathbb{R}$, the convex hull of the identified set for $F_{Y_1-Y_0}(z)$ is given by

$$\mathcal{I}_{\theta_{\text{DTE}}}^{\text{marg}}(F_{Y,X,W}; c) = \left[\mathbb{E} \left(\max \left\{ \sup_{y \in \mathbb{R}} (\underline{F}_{Y_1|X,W}(y | X, W) - \overline{F}_{Y_0|X,W}(y - z | X, W)), 0 \right\}, \right. \right. \\ \left. \left. 1 + \mathbb{E} \left(\min \left\{ \inf_{y \in \mathbb{R}} (\overline{F}_{Y_1|X,W}(y | X, W) - \underline{F}_{Y_0|X,W}(y - z | X, W)), 0 \right\} \right) \right) \right].$$

This expression involves two one-dimensional optimization problems, but the objective functions are known, closed-form functionals of the distribution of the observables. Bounds on the QDTE can be obtained as a corollary by taking the left-inverse of the cdf bounds.

6 Conclusion

In this paper we proposed a general class of relaxations of unconfoundedness, and showed how it includes several previous approaches as special cases. We then derived closed form identification results for many different target parameters under this general class of relaxations. There are at least three natural next steps. First, in this paper we focused on population level identification results. Corresponding estimation and inference results can likely be derived by using standard sample analog estimators and arguments, but we leave the details to future work. Second, it would be interesting to explore whether our bounds have either the double-sharpness or double-validity properties defined in Dorn et al. (2024), and if not, whether alternative bounds that had these properties could be derived. Third, it would be interesting to extend our results to independence assumptions beyond unconfoundedness, such as IV exogeneity (e.g., section 4 of Masten and Poirier 2018b).

References

- ASH, R. B. AND C. A. DOLÉANS-DADE (2000): *Probability and Measure Theory*, Academic Press.
- BILLINGSLEY, P. (1995): *Probability and Measure*, John Wiley & Sons, 3rd ed.

- BONVINI, M. AND E. H. KENNEDY (2022): “Sensitivity analysis via the proportion of unmeasured confounding,” *Journal of the American Statistical Association*, 117, 1540–1550.
- CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND Y. LUO (2018): “The sorted effects method: Discovering heterogeneous effects beyond their averages,” *Econometrica*, 86, 1911–1938.
- DING, P. AND T. J. VANDERWEELE (2016): “Sensitivity analysis without assumptions,” *Epidemiology*, 27, 368.
- DORN, J. AND K. GUO (2023): “Sharp sensitivity analysis for inverse propensity weighting via quantile balancing,” *Journal of the American Statistical Association*, 118, 2645–2657.
- DORN, J., K. GUO, AND N. KALLUS (2024): “Doubly-valid/doubly-sharp sensitivity analysis for causal inference with unmeasured confounding,” *Journal of the American Statistical Association*, 1–12.
- DORN, J. AND L. YAP (2024): “Sensitivity analysis for linear estimands,” *arXiv preprint arXiv:2309.06305*.
- DUARTE, G. (2024): “A unified approach for assessing sensitivity to violations of causal assumptions,” *Working paper*.
- FAN, Y. AND S. S. PARK (2010): “Sharp bounds on the distribution of treatment effects and their statistical inference,” *Econometric Theory*, 26, 931–951.
- GU, J., T. RUSSELL, AND T. STRINGHAM (2024): “Counterfactual identification and latent space enumeration in discrete outcome models,” *Review of Economic Studies (forthcoming)*.
- HECKMAN, J. J., J. SMITH, AND N. CLEMENTS (1997): “Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts,” *The Review of Economic Studies*, 64, 487–535.
- HUANG, M. AND S. D. PIMENTEL (2025): “Variance-based sensitivity analysis for weighting estimators results in more informative bounds,” *Biometrika*, 112.
- IMBENS, G. W. AND D. B. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge University Press.

- KALLUS, N. AND A. ZHOU (2018): “Confounding-robust policy improvement,” in *Advances in Neural Information Processing Systems*, vol. 31.
- MAKAROV, G. (1982): “Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed,” *Theory of Probability & its Applications*, 26, 803–806.
- MANSKI, C. F. (1997): “Monotone treatment response,” *Econometrica*, 1311–1334.
- (2003): *Partial Identification of Probability Distributions*, Springer.
- MASTEN, M. A. AND A. POIRIER (2016): “Partial independence in nonseparable models,” *cemmap working paper, CWP26/16*.
- (2018a): “Identification of treatment effects under conditional partial independence,” *Econometrica*, 86, 317–351.
- (2018b): “Salvaging falsified instrumental variable models,” *arXiv preprint arXiv:1812.11598v1*.
- (2020): “Inference on breakdown frontiers,” *Quantitative Economics*, 11, 41–111.
- (2023): “Choosing exogeneity assumptions in potential outcome models,” *The Econometrics Journal*, 26, 327–349.
- MASTEN, M. A., A. POIRIER, AND L. ZHANG (2024): “Assessing sensitivity to unconfoundedness: Estimation and inference,” *Journal of Business & Economic Statistics*, 42, 1–13.
- NELSEN, R. B. (2006): *An Introduction to Copulas*, Springer, second ed.
- RAMBACHAN, A., A. COSTON, AND E. KENNEDY (2023): “Robust design and evaluation of predictive algorithms under unobserved confounding,” *arXiv preprint arXiv:2212.09844*.
- ROSENBAUM, P. R. (1995): *Observational Studies*, Springer.
- (2002): *Observational Studies*, Springer, second ed.
- (2017): *Observation and Experiment: An Introduction to Causal Inference*, Harvard University Press.
- ROYDEN, H. AND P. M. FITZPATRICK (2010): *Real Analysis*, Pearson, 4th ed.

- SJÖLANDER, A. (2024): “Sharp bounds for causal effects based on Ding and VanderWeele’s sensitivity parameters,” *Journal of Causal Inference*, 12, 20230019.
- SKLAR, M. (1959): *Fonctions de répartition à n dimensions et leurs marges*, Université Paris 8.
- STOYE, J. (2010): “Partial identification of spread parameters,” *Quantitative Economics*, 1, 323–357.
- TAN, Z. (2006): “A distributional approach for causal inference using propensity scores,” *Journal of the American Statistical Association*, 101, 1619–1637.
- (2024): “Model-assisted sensitivity analysis for treatment effects under unmeasured confounding via regularized calibrated estimation,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86, 1339–1363.
- TORGOVITSKY, A. (2019): “Partial identification by extending subdistributions,” *Quantitative Economics*, 10, 105–144.
- VAN DER VAART, A. W. (2000): *Asymptotic Statistics*, Cambridge University Press.
- VANDERWEELE, T. J. AND P. DING (2017): “Sensitivity analysis in observational research: Introducing the E-value,” *Annals of Internal Medicine*, 167, 268–274.
- WILLIAMSON, R. C. AND T. DOWNS (1990): “Probabilistic arithmetic. I. Numerical methods for calculating convolutions and dependency bounds,” *International Journal of Approximate Reasoning*, 4, 89–158.
- ZHAO, Q., D. S. SMALL, AND B. B. BHATTACHARYA (2019): “Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 81, 735–761.

A Bound Expressions

In this section we summarize all the bounds that will be used in the proofs of lemmas and main results. These expressions are given for an arbitrary values of $(y, w) \in \mathbb{R} \times \text{supp}(W)$. Denote $\underline{c} := \underline{c}(w, \eta)$ and $\bar{c} := \bar{c}(w, \eta)$, where the dependence on w and η is implicitly understood.

A.1 Lower bound on Y_1

Let

$$\underline{\tau}_1 = \frac{(p_{1|w} - \underline{c})\bar{c}}{p_{1|w}(\bar{c} - \underline{c})} \quad \text{and} \quad \underline{Q}_1 = Q_{Y|X,W}(\underline{\tau}_1 | 1, w)$$

and

$$\begin{aligned} \underline{F}_{Y_1|W}(y | w) &= \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\bar{c}}, \frac{\underline{c} - p_{1|w}}{\underline{c}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}} \right\} \\ &= F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\bar{c}} \mathbb{1}(y < \underline{Q}_1) + \left(\frac{\underline{c} - p_{1|w}}{\underline{c}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}} \right) \mathbb{1}(y \geq \underline{Q}_1) \\ \underline{F}_{Y_1|X,W}(y | 0, w) &= \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1 - \bar{c})}{p_{0|w}\bar{c}}, \frac{\underline{c} - p_{1|w}}{\underline{c}p_{0|w}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1 - \underline{c})}{p_{0|w}\underline{c}} \right\} \\ &= F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1 - \bar{c})}{p_{0|w}\bar{c}} \mathbb{1}(y < \underline{Q}_1) + \left(\frac{\underline{c} - p_{1|w}}{\underline{c}p_{0|w}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1 - \underline{c})}{p_{0|w}\underline{c}} \right) \mathbb{1}(y \geq \underline{Q}_1) \end{aligned}$$

$$\underline{F}_{Y_1|X,W}(y | 1, w) = F_{Y|X,W}(y | 1, w).$$

Define

$$\underline{p}_1(y, w) = \bar{c} \mathbb{1}(y < \underline{Q}_1) + \underline{A}_1 \mathbb{1}(y = \underline{Q}_1) + \underline{c} \mathbb{1}(y > \underline{Q}_1)$$

where

$$\underline{A}_1 = \frac{\mathbb{P}(Y = \underline{Q}_1, X = 1 | W = w)}{\left(\frac{\underline{c} - p_{1|w}}{\underline{c}} + F_{Y|X,W}(\underline{Q}_1 | 1, w) \frac{p_{1|w}}{\underline{c}} \right) - \frac{p_{1|w}}{\bar{c}} \mathbb{P}(Y < \underline{Q}_1 | X = 1, W = w)}.$$

If the denominator is 0, set $\underline{A}_1 = p_{1|w}$. We can also derive the associated quantiles function bounds for $\tau \in (0, 1)$:

$$\begin{aligned} \bar{Q}_{Y_1|W}(\tau | w) &:= \underline{F}_{Y_1|W}^{-1}(\tau | w) = Q_{Y|X,W} \left(\min \left\{ \frac{\bar{c}}{p_{1|w}} \tau, \frac{p_{1|w} - \underline{c}}{p_{1|w}} + \frac{\underline{c}}{p_{1|w}} \tau \right\} | 1, w \right) \\ \bar{Q}_{Y_1|X,W}(\tau | 0, w) &:= \underline{F}_{Y_1|X}^{-1}(\tau | 0, w) = Q_{Y|X,W} \left(\min \left\{ \frac{\bar{c}p_{0|w}}{p_{1|w}(1 - \bar{c})} \tau, \frac{p_{1|w} - \underline{c}}{p_{1|w}(1 - \underline{c})} + \frac{\underline{c}p_{0|w}}{p_{1|w}(1 - \underline{c})} \tau \right\} | 1, w \right). \end{aligned}$$

A.2 Upper bound on Y_1

Let

$$\bar{\tau}_1 = 1 - \underline{\tau}_1 \quad \text{and} \quad \bar{Q}_1 = Q_{Y|X,W}(\bar{\tau}_1 | 1, w)$$

and

$$\begin{aligned}
\bar{F}_{Y_1|W}(y|w) &= \min \left\{ F_{Y|X,W}(y|1,w) \frac{p_{1|w}}{\underline{c}}, \frac{\bar{c} - p_{1|w}}{\bar{c}} + F_{Y|X,W}(y|1,w) \frac{p_{1|w}}{\bar{c}} \right\} \\
&= F_{Y|X,W}(y|1,w) \frac{p_{1|w}}{\underline{c}} \mathbb{1}(y < \bar{Q}_1) + \left(\frac{\bar{c} - p_{1|w}}{\bar{c}} + F_{Y|X,W}(y|1,w) \frac{p_{1|w}}{\bar{c}} \right) \mathbb{1}(y \geq \bar{Q}_1) \\
\bar{F}_{Y_1|X,W}(y|0,w) &= \min \left\{ F_{Y|X,W}(y|1,w) \frac{p_{1|w}(1-\underline{c})}{p_{0|w}\underline{c}}, \frac{\bar{c} - p_{1|w}}{\bar{c}p_{0|w}} + F_{Y|X,W}(y|1,w) \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}} \right\} \\
&= F_{Y|X,W}(y|1,w) \frac{p_{1|w}(1-\underline{c})}{(1-p_{1|w})\underline{c}} \mathbb{1}(y < \bar{Q}_1) + \left(\frac{\bar{c} - p_{1|w}}{\bar{c}p_{0|w}} + F_{Y|X,W}(y|1,w) \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}} \right) \mathbb{1}(y \geq \bar{Q}_1) \\
\bar{F}_{Y_1|X,W}(y|1,w) &= F_{Y|X,W}(y|1,w).
\end{aligned}$$

Define

$$\bar{p}_1(y,w) = \underline{c} \mathbb{1}(y < \bar{Q}_1) + \bar{A}_1 \mathbb{1}(y = \bar{Q}_1) + \bar{c} \mathbb{1}(y > \bar{Q}_1)$$

where

$$\bar{A}_1 = \frac{\mathbb{P}(Y = \bar{Q}_1, X = 1 | W = w)}{\left(\frac{\bar{c} - p_{1|w}}{\bar{c}} + F_{Y|X,W}(\bar{Q}_1 | 1, w) \frac{p_{1|w}}{\bar{c}} \right) - \frac{p_{1|w}}{\underline{c}} \mathbb{P}(Y < \bar{Q}_1 | X = 1, W = w)}.$$

If the denominator is 0, set $\bar{A}_1 = p_{1|w}$. We can also derive the associated quantiles function bounds for $\tau \in (0, 1)$:

$$\begin{aligned}
\underline{Q}_{Y_1|W}(\tau|w) &:= \bar{F}_{Y_1|W}^{-1}(\tau|w) = Q_{Y|X,W} \left(\max \left\{ \frac{\underline{c}}{p_{1|w}} \tau, \frac{p_{1|w} - \bar{c}}{p_{1|w}} + \frac{\bar{c}}{p_{1|w}} \tau \right\} | 1, w \right) \\
\underline{Q}_{Y_1|X,W}(\tau|0,w) &:= \bar{F}_{Y_1|X,W}^{-1}(\tau|0,w) = Q_{Y|X,W} \left(\max \left\{ \frac{\underline{c}p_{0|w}}{p_{1|w}(1-\underline{c})} \tau, \frac{p_{1|w} - \bar{c}}{p_{1|w}(1-\bar{c})} + \frac{\bar{c}p_{0|w}}{p_{1|w}(1-\bar{c})} \tau \right\} | 1, w \right).
\end{aligned}$$

A.3 Lower Bound on Y_0

Under c -dependence, we have that $\mathbb{P}(X = 0 | Y_0, W = w) = 1 - \mathbb{P}(X = 1 | Y_0, W = w) \in [1 - \bar{c}(w, \eta), 1 - \underline{c}(w, \eta)]$. So we can take the bound expressions for $F_{Y_1|W}$, swap $p_{1|w}$ and $p_{0|w}$, and swap (\underline{c}, \bar{c}) and $(1-\bar{c}, 1-\underline{c})$ and get the correct expressions for the bounds for $F_{Y_0|W}$. Here are all the expressions.

Let

$$\underline{\tau}_0 = \frac{(\bar{c} - p_{1|w})(1-\underline{c})}{p_{0|w}(\bar{c} - \underline{c})} \quad \text{and} \quad \underline{Q}_0 = Q_{Y|X,W}(\underline{\tau}_0|0, w)$$

and

$$\begin{aligned}
\underline{F}_{Y_0|W}(y | w) &= \max \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\underline{c}}, \frac{p_{1|w}-\bar{c}}{1-\bar{c}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\bar{c}} \right\} \\
&= F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\underline{c}} \mathbb{1}(y < \underline{Q}_0) + \left(\frac{p_{1|w}-\bar{c}}{1-\bar{c}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\bar{c}} \right) \mathbb{1}(y \geq \underline{Q}_0) \\
\underline{F}_{Y_0|X,W}(y | 1, w) &= \max \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1-\underline{c})}, \frac{p_{1|w}-\bar{c}}{(1-\bar{c})p_{1|w}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1-\bar{c})} \right\} \\
&= F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1-\underline{c})} \mathbb{1}(y < \underline{Q}_0) + \left(\frac{p_{1|w}-\bar{c}}{(1-\bar{c})p_{1|w}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1-\bar{c})} \right) \mathbb{1}(y \geq \underline{Q}_0) \\
\underline{F}_{Y_0|X,W}(y | 0, w) &= F_{Y|X}(y | 0, w).
\end{aligned}$$

Define

$$\underline{p}_0(y, w) = \underline{c} \mathbb{1}(y < \underline{Q}_0) + (1 - \underline{A}_0) \mathbb{1}(y = \underline{Q}_0) + \bar{c} \mathbb{1}(y > \underline{Q}_0)$$

where

$$\underline{A}_0 = \frac{p_{0|w}(F_{Y|X,W}(\underline{Q}_0 | 0, w) - F_{Y|X,W}(\underline{Q}_0 - | 0, w))}{\underline{F}_{Y_0|W}(\underline{Q}_0 | w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - | w)}.$$

If the denominator is 0, set $\underline{A}_0 = p_{0|w}$. We can also derive the associated quantiles function bounds for $\tau \in (0, 1)$:

$$\begin{aligned}
\bar{Q}_{Y_0|W}(\tau | w) &:= \underline{F}_{Y_0|W}^{-1}(\tau) = Q_{Y|X,W} \left(\min \left\{ \frac{1-\underline{c}}{p_{0|w}} \tau, \frac{p_{0|w} - (1-\bar{c})}{p_{0|w}} + \frac{1-\bar{c}}{p_{0|w}} \tau \right\} | 0, w \right) \\
\bar{Q}_{Y_0|X,W}(\tau | 1, w) &:= \underline{F}_{Y_0|X}^{-1}(\tau | 1) = Q_{Y|X,W} \left(\min \left\{ \frac{(1-\underline{c})p_{1|w}}{p_{0|w}\underline{c}} \tau, \frac{p_{0|w} - (1-\bar{c})}{p_{0|w}\bar{c}} + \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}} \tau \right\} | 0, w \right).
\end{aligned}$$

A.4 Upper bound for Y_0

Let

$$\bar{\tau}_0 = 1 - \underline{\tau}_0 \quad \text{and} \quad \bar{Q}_0 = Q_{Y|X,W}(\bar{\tau}_0 | 0, w)$$

and

$$\begin{aligned}
\bar{F}_{Y_0|W}(y | w) &= \min \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\bar{c}}, \frac{p_{1|w}-\underline{c}}{1-\underline{c}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\underline{c}} \right\} \\
&= F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\bar{c}} \mathbb{1}(y < \bar{Q}_0) + \left(\frac{p_{1|w}-\underline{c}}{1-\underline{c}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1-\underline{c}} \right) \mathbb{1}(y \geq \bar{Q}_0) \\
\bar{F}_{Y_0|X,W}(y | 1, w) &= \min \left\{ F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1-\bar{c})}, \frac{p_{1|w}-\underline{c}}{(1-\underline{c})p_{1|w}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1-\underline{c})} \right\} \\
&= F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1-\bar{c})} \mathbb{1}(y < \bar{Q}_0) + \left(\frac{p_{1|w}-\underline{c}}{(1-\underline{c})p_{1|w}} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1-\underline{c})} \right) \mathbb{1}(y \geq \bar{Q}_0).
\end{aligned}$$

Define

$$\bar{p}_0(y, w) = \bar{c} \mathbb{1}(y < \bar{Q}_0) + (1 - \bar{A}_0) \mathbb{1}(y = \bar{Q}_0) + \underline{c} \mathbb{1}(y > \bar{Q}_0)$$

where

$$\bar{A}_0 = \frac{\mathbb{P}(Y = \bar{Q}_0, X = 0 \mid W = w)}{\left(\frac{p_{1|w} - \underline{c}}{1 - \underline{c}} + F_{Y|X,W}(\bar{Q}_0|0, w) \frac{p_{0|w}}{1 - \underline{c}} \right) - \frac{p_{0|w}}{1 - \bar{c}} \mathbb{P}(Y < \bar{Q}_0 \mid X = 0, W = w)}.$$

If the denominator is 0, set $\bar{A}_0 = p_{0|w}$. We can also derive the associated quantiles function bounds for $\tau \in (0, 1)$:

$$\underline{Q}_{Y_0|W}(\tau \mid w) := \bar{F}_{Y_0|W}^{-1}(\tau \mid w) = Q_{Y|X,W} \left(\max \left\{ \frac{1 - \bar{c}}{p_{0|w}} \tau, \frac{p_{0|w} - (1 - \underline{c})}{p_{0|w}} + \frac{1 - \underline{c}}{p_{0|w}} \tau \right\} \mid 0, w \right)$$

$$\underline{Q}_{Y_0|X,W}(\tau \mid 1, w) := \bar{F}_{Y_0|X,W}^{-1}(\tau \mid 1, w) = Q_{Y|X,W} \left(\max \left\{ \frac{(1 - \bar{c})p_{1|w}}{p_{0|w}\bar{c}} \tau, \frac{p_{0|w} - (1 - \underline{c})}{p_{0|w}\underline{c}} + \frac{(1 - \underline{c})p_{1|w}}{p_{0|w}\underline{c}} \tau \right\} \mid 0, w \right).$$

B Proofs for Section 3

Proof of Lemma 1. Let $x \in \{0, 1\}$ and fix $w \in \text{supp}(W)$. By the law of iterated expectations, $\mathbb{E}[X \mid Y_x, W = w] = \mathbb{E}[\mathbb{E}[X \mid Y_1, Y_0, W = w] \mid Y_x, W = w]$. Since $\mathbb{E}[X \mid Y_1, Y_0, W = w] \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely, we then have that $\mathbb{E}[X \mid Y_x, W = w] \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely as well. \square

Proof of Proposition 1. Part 1: Suppose marginal c -dependence holds with bound functions $[\underline{c}(w, \eta), \bar{c}(w, \eta)]$. Fix $(x, w) \in \{0, 1\} \times \text{supp}(W)$. We have that

$$R_x(Y_x, w) = \frac{p_x(Y_x, w)}{1 - p_x(Y_x, w)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}} \in \left[\frac{\underline{c}(w, \eta)}{1 - \underline{c}(w, \eta)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}}, \frac{\bar{c}(w, \eta)}{1 - \bar{c}(w, \eta)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}} \right], \quad (8)$$

where the inclusion holds from the mapping $a \mapsto a/(1 - a)$ being strictly increasing over $a \in (0, 1)$ and from $p_x(Y_x, w) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)] \subset (0, 1)$ almost surely. We note that

$$\frac{\bar{c}(w, \eta)}{1 - \bar{c}(w, \eta)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}} \in \left[\frac{p_{1|w}}{1 - p_{1|w}} \Big/ \frac{p_{1|w}}{1 - p_{1|w}}, +\infty \right) = [1, +\infty)$$

where the inclusion holds from $\bar{c}(w, \eta) \in [p_{1|w}, 1)$. Similarly,

$$\frac{\underline{c}(w, \eta)}{1 - \underline{c}(w, \eta)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}} \in \left(\frac{0}{1 - 0} \Big/ \frac{p_{1|w}}{1 - p_{1|w}}, \frac{p_{1|w}}{1 - p_{1|w}} \Big/ \frac{p_{1|w}}{1 - p_{1|w}} \right] = (0, 1]$$

from $\underline{c}(w, \eta) \in (0, p_{1|w}]$. We conclude that the GSM holds with the bound functions from equation (4). Replacing marginal c -dependence with joint c -dependence delivers the same bounds for the GJSM.

Part 2: Suppose the GSM holds with bound functions $[\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)]$. Fix $(x, w) \in \{0, 1\} \times \text{supp}(W)$. We have that

$$p_x(Y_x, w) = \frac{p_{1|w} R_x(Y_x, w)}{p_{0|w} + p_{1|w} R_x(Y_x, w)} \in \left[\frac{p_{1|w} \underline{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w} \underline{\Lambda}(w, \eta)}, \frac{p_{1|w} \bar{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w} \bar{\Lambda}(w, \eta)} \right]. \quad (9)$$

The equality holds from inverting the equation $R_x(Y_x, w) = \frac{p_x(Y_x, w)}{1 - p_x(Y_x, w)} \Big/ \frac{p_{1|w}}{1 - p_{1|w}}$ in $p_x(Y_x, w)$. The inclusion holds from the mapping $a \mapsto a/(1 + a)$ being strictly increasing for $a \in [0, +\infty)$ and from $R_x(Y_x, w) \in [\underline{\Lambda}(w, \eta), \bar{\Lambda}(w, \eta)] \subset (0, +\infty)$ almost surely. We note that

$$\frac{p_{1|w} \underline{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w} \underline{\Lambda}(w, \eta)} \in \left(\frac{p_{1|w} \cdot 0}{p_{0|w} + p_{1|w} \cdot 0}, \frac{p_{1|w} \cdot 1}{p_{0|w} + p_{1|w} \cdot 1} \right] = (0, p_{1|w}]$$

by $\underline{\Lambda}(w, \eta) \in (0, 1]$. Similarly,

$$\frac{p_{1|w}\bar{\Lambda}(w, \eta)}{p_{0|w} + p_{1|w}\bar{\Lambda}(w, \eta)} \in \left[\frac{p_{1|w} \cdot 1}{p_{0|w} + p_{1|w} \cdot 1}, 1 \right) = [p_{1|w}, 1)$$

by $\bar{\Lambda}(w, \eta) \in [1, +\infty]$. We conclude that marginal c -dependence holds with the bound functions from equation (1). Replacing the GSM with the GJSM delivers the same bounds for joint c -dependence. \square

C Proofs for Section 4

Proof of Lemma 2. By Lemma 1, it suffices to show the desired results under Assumption 2. Let $y \in \mathbb{R}$ and $w \in \text{supp}(W)$ be fixed. Note that

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} &= \mathbb{E} \left[\frac{\mathbb{1}(Y_1 \leq y)X}{p_1(Y_1, w)} \mid W = w \right] \\ &= \mathbb{E} \left[\frac{\mathbb{1}(Y_1 \leq y)\mathbb{E}[X \mid Y_1, W = w]}{p_1(Y_1, w)} \mid W = w \right] \\ &= \mathbb{E} [\mathbb{1}(Y_1 \leq y) \mid W = w] \\ &= F_{Y_1|W}(y \mid w), \end{aligned}$$

where the second equality follows from the law of iterated expectations and the third from $p_1(Y_1, w) \geq \underline{c}(w, \eta) > 0$ almost surely by Assumption 2. Likewise, we have

$$\mathbb{E} \left[\frac{\mathbb{1}(Y > y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} = 1 - F_{Y_1|W}(y \mid w).$$

Therefore,

$$\begin{aligned} F_{Y_1|W}(y \mid w) &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} \\ &\leq \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{\underline{c}(w, \eta)} \mid X = 1, W = w \right] p_{1|w} \\ &= F_{Y|X, W}(y \mid 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)} \end{aligned}$$

and

$$\begin{aligned} F_{Y_1|W}(y \mid w) &= 1 - \mathbb{P}(Y_1 > y \mid W = w) \\ &= 1 - \mathbb{E} \left[\frac{\mathbb{1}(Y > y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} \\ &\leq 1 - \mathbb{E} \left[\frac{\mathbb{1}(Y > y)}{\bar{c}(w, \eta)} \mid X = 1, W = w \right] p_{1|w} \\ &= \frac{\bar{c}(w, \eta) - p_{1|w}}{\bar{c}(w, \eta)} + F_{Y|X, W}(y \mid 1, w) \frac{p_{1|w}}{\bar{c}(w, \eta)}. \end{aligned}$$

The inequalities follow from $p_1(Y_1, w)^{-1} \in [\bar{c}(w, \eta)^{-1}, \underline{c}(w, \eta)^{-1}]$ almost surely. By these two inequalities,

$$F_{Y_1|W}(y \mid w) \leq \min \left\{ F_{Y|X, W}(y \mid 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)}, \frac{\bar{c}(w, \eta) - p_{1|w}}{\bar{c}(w, \eta)} + F_{Y|X, W}(y \mid 1, w) \frac{p_{1|w}}{\bar{c}(w, \eta)} \right\} = \bar{F}_{Y_1|W}(y \mid w).$$

Similarly,

$$\begin{aligned}
F_{Y_1|W}(y | w) &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} \\
&\geq \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{\bar{c}(w, \eta)} \mid X = 1, W = w \right] p_{1|w} \\
&= F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\bar{c}(w, \eta)}
\end{aligned}$$

and

$$\begin{aligned}
F_{Y_1|W}(y | w) &= 1 - \mathbb{P}(Y_1 > y \mid W = w) \\
&= 1 - \mathbb{E} \left[\frac{\mathbb{1}(Y > y)}{p_1(Y, w)} \mid X = 1, W = w \right] p_{1|w} \\
&\geq 1 - \mathbb{E} \left[\frac{\mathbb{1}(Y > y)}{\underline{c}(w, \eta)} \mid X = 1, W = w \right] p_{1|w} \\
&= \frac{\underline{c}(w, \eta) - p_{1|w}}{\underline{c}(w, \eta)} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)}.
\end{aligned}$$

By these two inequalities,

$$F_{Y_1|W}(y | w) \geq \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\bar{c}(w, \eta)}, \frac{\underline{c}(w, \eta) - p_{1|w}}{\underline{c}(w, \eta)} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}}{\underline{c}(w, \eta)} \right\} = \underline{F}_{Y_1|W}(y | w).$$

Therefore we have established $F_{Y_1|W}(y | w) \in [\underline{F}_{Y_1|W}(y | w), \bar{F}_{Y_1|W}(y | w)]$, as desired.

Next we establish the bounds for $F_{Y_0|W}(y | w)$. We also have that

$$\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)}{1 - p_0(Y, w)} \mid X = 0, W = w \right] p_{0|w} = F_{Y_0|W}(y | w).$$

Furthermore, note that $\mathbb{E}[(1 - X) \mid Y_0, W = w] = \mathbb{P}(X = 0 \mid Y_0, W = w) \in [1 - \bar{c}(w, \eta), 1 - \underline{c}(w, \eta)]$. Changing $X = 1$ to $X = 0$, and $(\underline{c}(w, \eta), \bar{c}(w, \eta))$ to $(1 - \bar{c}(w, \eta), 1 - \underline{c}(w, \eta))$ yields

$$\begin{aligned}
F_{Y_0|W}(y | w) &\leq F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \bar{c}(w, \eta)} \\
F_{Y_0|W}(y | w) &\leq \frac{p_{1|w} - \underline{c}(w, \eta)}{1 - \underline{c}(w, \eta)} + F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \underline{c}(w, \eta)} \\
F_{Y_0|W}(y | w) &\geq F_{Y|X,W}(y | 0, w) \frac{p_{0|w}}{1 - \underline{c}(w, \eta)} \\
F_{Y_0|W}(y | w) &\geq \frac{p_{1|w} - \bar{c}(w, \eta)}{1 - \bar{c}(w, \eta)} + F_{Y|X,W}(y | 0) \frac{p_{0|w}}{1 - \bar{c}(w, \eta)}.
\end{aligned}$$

almost surely. Therefore, $F_{Y_0|W}(y | w) \in [\underline{F}_{Y_0|W}(y | w), \bar{F}_{Y_0|W}(y | w)]$. \square

C.1 Proof of Theorem 1

We provide and show a number of preliminary lemmas that are used to prove Theorem 1. This first lemma establishes some properties of cdf bounds for Y_x given (X, W) .

Lemma 4 (Bounds of CDFs). Let assumptions 1 and 2 hold. Then, for $x \in \{0, 1\}$ and $w \in \text{supp}(W)$,

1. The functions $\underline{F}_{Y_x|X,w}(\cdot | 1 - x, w)$ and $\bar{F}_{Y_x|X,W}(\cdot | 1 - x, w)$, which are defined in Appendix A, are cdfs;

2. For all $y \in \mathbb{R}$,

$$\begin{aligned}\underline{F}_{Y_x|X,W}(y | 1-x, w)p_{1-x|w} + F_{Y|X,W}(Y | X, w)p_{x|w} &= \underline{F}_{Y_x|W}(y | w) \\ \overline{F}_{Y_x|X,W}(y | 1-x, w)p_{1-x|w} + F_{Y|X,W}(Y | X, w)p_{x|w} &= \overline{F}_{Y_x|W}(y | w).\end{aligned}$$

Proof of Lemma 4. Proof of Part 1: We show that $\underline{F}_{Y_1|X,W}(y | 0, w)$ is a cdf by showing it is nondecreasing, has limits $(0, 1)$ when y approaches $(-\infty, +\infty)$, and is right-continuous. The same arguments can be used to deduce that $\overline{F}_{Y_1|X,W}(y | 0, w)$, $\underline{F}_{Y_0|X,W}(y | 1, w)$, and $\overline{F}_{Y_0|X,W}(y | 1, w)$ are also cdfs.

The function

$$\underline{F}_{Y_1|X,W}(y | 0, w) = \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}}, \frac{\underline{c} - p_{1|w}}{\underline{c}p_{0|w}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1-\underline{c})}{\underline{c}p_{0|w}} \right\}$$

is nondecreasing in y since each of its two arguments is nondecreasing in y , due to $F_{Y|X,W}(\cdot | 1, w)$ being a cdf. Then note that

$$\lim_{y \rightarrow \infty} \underline{F}_{Y_1|X,W}(y | 0, w) = \max \left\{ \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}}, 1 \right\} = \max \left\{ \frac{p_{1|w} - p_{1|w}\bar{c}}{\bar{c} - p_{1|w}\bar{c}}, 1 \right\} = 1,$$

where the last equality follows by $p_{1|w} \leq \bar{c}$. Also note that

$$\lim_{y \rightarrow -\infty} \underline{F}_{Y_1|X,W}(y | 0, w) = \max \left\{ 0, \frac{\underline{c} - p_{1|w}}{\underline{c}p_{0|w}} \right\} = 0,$$

where the last equality follows by $\underline{c} \leq p_{1|w}$. Finally, we can see that $\underline{F}_{Y_1|X,W}(y | 0, w)$ is right-continuous with respect to y since $F_{Y|X,W}(y | 1, w)$ is right-continuous and by the continuity of affine transformations and of the maximum function. Therefore, $\underline{F}_{Y_1|X,W}(y | 0, w)$ is a cdf.

Proof of Part 2: We show the first equality with $x = 1$, and the same arguments can be used to establish the equality for other cases. For $y \in \mathbb{R}$, the desired result follows by the following derivations:

$$\begin{aligned}\underline{F}_{Y_1|X,W}(y | 0, w)p_{0|w} + F_{Y|X,W}(y | 1, w)p_{1|w} \\ &= \max \left\{ F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1-\bar{c})}{p_{0|w}\bar{c}}, \frac{\underline{c} - p_{1|w}}{\underline{c}p_{0|w}} + F_{Y|X,W}(y | 1, w) \frac{p_{1|w}(1-\underline{c})}{p_{0|w}\underline{c}} \right\} p_{0|w} + F_{Y|X,W}(y | 1, w)p_{1|w} \\ &= \max \left\{ F_{Y|X,W}(y | 1, w)p_{1|w} \left(\frac{1-\bar{c}}{\bar{c}} + 1 \right), \frac{\underline{c} - p_{1|w}}{\underline{c}} + F_{Y|X,W}(y | 1, w)p_{1|w} \left(\frac{1-\underline{c}}{\underline{c}} + 1 \right) \right\} \\ &= \max \left\{ \frac{F_{Y|X,W}(y | 1, w)p_{1|w}}{\bar{c}}, \frac{\underline{c} - p_{1|w}}{\underline{c}} + \frac{F_{Y|X,W}(y | 1, w)p_{1|w}}{\underline{c}} \right\} \\ &= \underline{F}_{Y_1|W}(y | w).\end{aligned}$$

Thus the proof is complete. \square

Lemma 5. Let $x \in \{0, 1\}$ and $w \in \text{supp}(W)$. Suppose $m(\cdot)$ is a Borel measurable function and $\mathbb{P}(m(Y_x) \geq \delta | W = w) = 1$ for some $\delta > 0$. The following statements are equivalent:

1. Conditional on $W = w$, the following statement holds almost surely:

$$m(Y_x) = \mathbb{P}(X = x | Y_x, W = w). \quad (10)$$

2. For all $y \in \mathbb{R}$, the following equality holds:

$$\mathbb{E} \left[\frac{\mathbb{1}(Y_x \leq y) \mathbb{1}(X = x)}{m(Y_x)} | W = w \right] = \mathbb{P}(Y_x \leq y | W = w). \quad (11)$$

Proof of Lemma 5. We first prove that (10) implies (11). This follows from the law of iterated expectations:

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{1}(Y_x \leq y) \mathbb{1}(X = x)}{m(Y_x)} \mid W = w \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y_x \leq y) \mathbb{E}[\mathbb{1}(X = x) \mid Y_x, W = w]}{m(Y_x)} \mid W = w \right] \\ &= \mathbb{E} \left[\frac{\mathbb{1}(Y_x \leq y) \mathbb{P}(X = x \mid Y_x, W = w)}{m(Y_x)} \mid W = w \right] \\ &= \mathbb{E}[\mathbb{1}(Y_x \leq y) \mid W = w] \\ &= \mathbb{P}(Y_x \leq y \mid W = w), \end{aligned}$$

where we use (10) and the assumption $m(Y_x) \geq \delta > 0$ for the third equality.

Next, we prove that (11) implies (10). To show this result, we first establish a few key facts:

1. By the law of iterated expectations, (11) implies

$$\mathbb{E} \left[\mathbb{1}(Y_x \leq y) \frac{\mathbb{P}(X = x \mid Y_x, W = w) - m(Y_x)}{m(Y_x)} \mid W = w \right] = 0$$

for all $y \in \mathbb{R}^2$.

2. For $y \in \mathbb{R}$, define the preimage from a half-space on \mathbb{R}^2 as

$$I_y = \{\omega \in \Omega : Y_x(\omega) \leq y\},$$

where Ω denotes Y_x 's sample space. Let $\mathcal{A} = \{I_y : y \in \mathbb{R}\}$. We then note that \mathcal{A} is closed under intersection since

$$I_y \cap I_{y'} = I_{\min\{y, y'\}} \in \mathcal{A} \quad \text{for any } y, y' \in \mathbb{R}.$$

This, combined with the non-emptiness of \mathcal{A} , implies that \mathcal{A} is a π -system.

3. The sample space can be written as a countable union of sets in \mathcal{A} since

$$\Omega = \{\omega \in \Omega : Y_x(\omega) < \infty\} = \bigcup_{n=1}^{\infty} I_n.$$

4. The random variable $[\mathbb{P}(X = x \mid Y_x, W = w) - m(Y_x)]/m(Y_x)$ is measurable with respect to the σ -algebra generated by Y_x due to the Borel measurability of $m(\cdot)$, and it is integrable since

$$\left| \frac{\mathbb{P}(X = x \mid Y_x, W = w) - m(Y_x)}{m(Y_x)} \right| \leq \frac{\mathbb{P}(X = x \mid Y_x, W = w)}{m(Y_x)} + 1 \leq \frac{1}{\delta} + 1 < +\infty,$$

where the first inequality follows by triangle inequality, and the second inequality follows by the assumption that $m(Y_x) \geq \delta > 0$ almost surely.

Given the above facts, it follows by Billingsley (1995, Theorem 34.1) that

$$\frac{\mathbb{P}(X = x \mid Y_x, W = w) - m(Y_x)}{m(Y_x)} = 0 \quad \text{with probability one conditional on } W = w.$$

From this equality we conclude that $\mathbb{P}(X = x \mid Y_x, W = w) = m(Y_x)$ with probability one conditional on $W = w$ since $m(Y_x) \geq \delta > 0$ almost surely. So the desired result has been established. \square

The following lemma is a subset of Lemma 21.1 in van der Vaart (2000), so we omit its proof.

Lemma 6 (Properties of CDFs and Quantiles). Let $p \in (0, 1)$ and $x \in \mathbb{R}$. Let F be a cdf and $Q(p) = \inf\{z \in \mathbb{R} : F(z) \geq p\}$ be its quantile function. Then,

1. $Q(p) \leq x$ if and only if $p \leq F(x)$;
2. $F(Q(p)) \geq p$ where equality can fail only if F is discontinuous at $Q(p)$;

3. $F(Q(p)-) \leq p$.

This next lemma is a compendium of properties of the cdf bounds. Its results are used throughout our proofs for the main theorems.

Lemma 7 (Preliminary Results). Let $w \in \text{supp}(W)$ and suppose assumptions 1 and 2 hold with $\underline{c}(w, \eta) < p_{1|w} < \bar{c}(w, \eta)$. Then, for $x \in \{0, 1\}$,

1. $\underline{\tau}_x, \bar{\tau}_x \in (0, 1)$;
2. $\bar{F}_{Y_x|W}(y | w)$ is continuous at $y = \bar{Q}_x$ if and only if $\mathbb{P}(Y = \bar{Q}_x | X = x, W = w) = 0$, and $\underline{F}_{Y_x|W}(y | w)$ is continuous at $y = \underline{Q}_x$ if and only if $\mathbb{P}(Y = \underline{Q}_x | X = x, W = w) = 0$;
3. $\bar{A}_1, \underline{A}_1, 1 - \bar{A}_0, 1 - \underline{A}_0 \in [\underline{c}, \bar{c}]$;
4. $\bar{p}_x(Y_x, w), \underline{p}_x(Y_x, w) \in [\underline{c}, \bar{c}]$ almost surely;
5. For all $y \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{\underline{p}_1(Y, w)} | W = w \right] &= \underline{F}_{Y_1|W}(y | w) \quad \text{and} \quad \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{\bar{p}_1(Y, w)} | W = w \right] = \bar{F}_{Y_1|W}(y | w), \\ \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - \underline{p}_0(Y, w)} | W = w \right] &= \underline{F}_{Y_0|W}(y | w) \quad \text{and} \quad \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - \bar{p}_0(Y, w)} | W = w \right] = \bar{F}_{Y_0|W}(y | w). \end{aligned}$$

6. For all $(y_1, y_0) \in \mathbb{R}^2$, the following inequalities are equivalent:

- (a) $\bar{F}_{Y_x|W}(y_x|w) \leq \underline{F}_{Y_{1-x}|W}(y_{1-x}|w)$;
- (b) $F_{Y|X,W}(y_x|x, w) \leq \underline{F}_{Y_{1-x}|X,W}(y_{1-x}|x, w)$;
- (c) $\bar{F}_{Y_x|X,W}(y_x | 1 - x, w) \leq F_{Y|X,W}(y_{1-x} | 1 - x, w)$.

Also, the following inequalities are equivalent:

- (d) $\bar{F}_{Y_x|W}(y_x|w) \geq \underline{F}_{Y_{1-x}|W}(y_{1-x}|w)$;
- (e) $F_{Y|X,W}(y_x|x, w) \geq \underline{F}_{Y_{1-x}|X,W}(y_{1-x}|x, w)$;
- (f) $\bar{F}_{Y_x|X,W}(y_x | 1 - x, w) \geq F_{Y|X,W}(y_{1-x} | 1 - x, w)$.

7. The following inequalities are equivalent:

- (a) $F_{Y|X,W}(\bar{Q}_1 - | 1, w) \leq \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w)$;
- (b) $\bar{F}_{Y_1|X,W}(\bar{Q}_1 - | 0, w) \leq F_{Y|X,W}(\underline{Q}_0 - | 0, w)$;
- (c) $\bar{F}_{Y_1|W}(\bar{Q}_1 - | w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 - | w)$.

Also, the following inequalities are equivalent:

- (d) $F_{Y|X,W}(\underline{Q}_1 - | 1, w) \leq \bar{F}_{Y_0|X,W}(\bar{Q}_0 - | 1, w)$;
- (e) $\underline{F}_{Y_1|X,W}(\underline{Q}_1 - | 0, w) \leq F_{Y|X,W}(\bar{Q}_0 - | 0, w)$;
- (f) $\underline{F}_{Y_1|W}(\underline{Q}_1 - | w) \leq \bar{F}_{Y_0|W}(\bar{Q}_0 - | w)$.

8. The following inequalities hold:

$$\max \left\{ \bar{F}_{Y_1|W}(\bar{Q}_1 - | w), \underline{F}_{Y_0|W}(\underline{Q}_0 - | w) \right\} \leq \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \leq \min \left\{ \bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w) \right\}$$

and

$$\max \left\{ \bar{F}_{Y_0|W}(\bar{Q}_0 - | w), \underline{F}_{Y_1|W}(\underline{Q}_1 - | w) \right\} \leq \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} \leq \min \left\{ \bar{F}_{Y_0|W}(\bar{Q}_0 | w), \underline{F}_{Y_1|W}(\underline{Q}_1 | w) \right\}.$$

9. The following inequalities hold:

$$\max \left\{ F_{Y|X,W}(\bar{Q}_1 - | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w) \right\} \leq \bar{\tau}_1 \leq \min \left\{ F_{Y|X,W}(\bar{Q}_1 | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w) \right\}$$

and

$$\max \left\{ \bar{F}_{Y_0|X,W}(\bar{Q}_0 - | 1, w), F_{Y|X,W}(\underline{Q}_1 - | 1, w) \right\} \leq \underline{\tau}_1 \leq \min \left\{ \bar{F}_{Y_0|X,W}(\bar{Q}_0 | 1, w), F_{Y|X,W}(\underline{Q}_1 | 1, w) \right\}.$$

10. For all $(y_1, y_0) \in \mathbb{R}^2$, the following inequalities are equivalent:

- (a) $\bar{F}_{Y_x|W}(y_x|w) + \bar{F}_{Y_{1-x}|W}(y_{1-x}|w) \geq 1$;
- (b) $F_{Y|X,W}(y_x|x, w) + \bar{F}_{Y_{1-x}|X,W}(y_{1-x}|x, w) \geq 1$;
- (c) $\bar{F}_{Y_x|X,W}(y_x | 1 - x, w) + F_{Y|X,W}(y_{1-x} | 1 - x, w) \geq 1$.

Also, the following inequalities are equivalent:

- (d) $\underline{F}_{Y_x|W}(y_x|w) + \underline{F}_{Y_{1-x}|W}(y_{1-x}|w) \geq 1$;
- (e) $F_{Y|X,W}(y_x|x, w) + \underline{F}_{Y_{1-x}|X,W}(y_{1-x}|x, w) \geq 1$;
- (f) $\underline{F}_{Y_x|X,W}(y_x | 1 - x, w) + F_{Y|X,W}(y_{1-x} | 1 - x, w) \geq 1$.

11. The following inequalities are equivalent:

- (a) $\bar{F}_{Y_1|W}(\bar{Q}_1 | w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - | w) \geq 1$;
- (b) $F_{Y|X,W}(\bar{Q}_1 | 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 - | 1, w) \geq 1$;
- (c) $\bar{F}_{Y_1|X,W}(\bar{Q}_1|0, w) + F_{Y|X,W}(\bar{Q}_0 - |0, w) \geq 1$.

Also, the following inequalities are equivalent:

- (d) $\bar{F}_{Y_1|W}(\bar{Q}_1 - | w) + \bar{F}_{Y_0|W}(\bar{Q}_0|w) \geq 1$;
- (e) $F_{Y|X,W}(\bar{Q}_1 - | 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 | 1, w) \geq 1$;
- (f) $\bar{F}_{Y_1|X,W}(\bar{Q}_1 - |0, w) + F_{Y|X,W}(\bar{Q}_0|0, w) \geq 1$.

Proof of Lemma 7. For brevity, we omit covariates $w \in \text{supp}(W)$ and drop notation referring on conditional probability $\cdot | w$ and $\cdot | W = w$ from this proof. However, note that our arguments hold when conditioning on $W = w$ throughout.

Proof of Part 1: First consider $\underline{\tau}_1$:

$$\underline{\tau}_1 = \frac{p_1 - \underline{c}}{\bar{c} - \underline{c}} \frac{\bar{c}}{p_1} = \frac{p_1 \bar{c} - \underline{c} \bar{c}}{p_1 \bar{c} - \underline{c} p_1} < \frac{p_1 \bar{c} - \underline{c} \bar{c}}{p_1 \bar{c} - \underline{c} \bar{c}} = 1,$$

where the inequality is strict because $p_1 < \bar{c}$ and $\underline{c} > 0$. Similarly,

$$\bar{\tau}_1 = \frac{p_1 - \underline{c}}{\bar{c} - \underline{c}} \frac{\bar{c}}{p_1} > \frac{\underline{c} - \underline{c}}{\bar{c} - \underline{c}} \frac{\bar{c}}{p_1} = 0$$

where the inequality is strict because $p_1 > \underline{c}$ and $\bar{c} > 0$. Thus, $\underline{\tau}_1 \in (0, 1)$. Since $\bar{\tau}_1 = 1 - \underline{\tau}_1$, $\bar{\tau}_1 \in (0, 1)$ as well. The proofs for $\bar{\tau}_0$ and $\underline{\tau}_0$ are similar.

Proof of Part 2: First consider the statement involving \bar{Q}_x with $x = 1$. We show the following inequality

$$\frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 | X = 1) \leq \bar{F}_{Y_1}(\bar{Q}_1) - \bar{F}_{Y_1}(\bar{Q}_1 -) \leq \frac{p_1}{\underline{c}} \mathbb{P}(Y = \bar{Q}_1 | X = 1). \quad (12)$$

From this inequality and Assumption 2 that $0 < \underline{c} \leq p_1 \leq \bar{c}$, we will conclude that $\bar{F}_{Y_1}(y)$ is continuous at $y = \bar{Q}_1$ if and only if $\mathbb{P}(Y = \bar{Q}_1 \mid X = 1) = 0$.

To show the lower bound inequality in (12), note that

$$\begin{aligned} \bar{F}_{Y_1}(\bar{Q}_1) - \bar{F}_{Y_1}(\bar{Q}_1-) &= \frac{\bar{c} - p_1}{\bar{c}} + F_{Y|X}(\bar{Q}_1 \mid 1) \frac{p_1}{\bar{c}} - \frac{p_1}{\bar{c}} F_{Y|X}(\bar{Q}_1 - \mid 1) \\ &= \frac{\bar{c} - p_1}{\bar{c}} - p_1 F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 \mid 1) - \mid 1) \frac{\bar{c} - \underline{c}}{\bar{c}\underline{c}} + \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1) \\ &\geq \frac{\bar{c} - p_1}{\bar{c}} - p_1 \bar{\tau}_1 \frac{\bar{c} - \underline{c}}{\bar{c}\underline{c}} + \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1) \\ &= \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1), \end{aligned}$$

The first line holds by the definition of \bar{F}_{Y_1} . The third line holds by Lemma 6.3. The last line holds by the definition of $\bar{\tau}_1$.

Likewise, we also have the following derivation:

$$\begin{aligned} \bar{F}_{Y_1}(\bar{Q}_1) - \bar{F}_{Y_1}(\bar{Q}_1-) &= \frac{\bar{c} - p_1}{\bar{c}} + F_{Y|X}(\bar{Q}_1 \mid 1) \frac{p_1}{\bar{c}} - \frac{p_1}{\bar{c}} F_{Y|X}(\bar{Q}_1 - \mid 1) \\ &= \frac{\bar{c} - p_1}{\bar{c}} - p_1 F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 \mid 1) \mid 1) \frac{\bar{c} - \underline{c}}{\bar{c}\underline{c}} + \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1) \\ &\leq \frac{\bar{c} - p_1}{\bar{c}} - p_1 \bar{\tau}_1 \frac{\bar{c} - \underline{c}}{\bar{c}\underline{c}} + \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1) \\ &= \frac{p_1}{\bar{c}} \mathbb{P}(Y = \bar{Q}_1 \mid X = 1), \end{aligned}$$

where we use Lemma 6.2 in the third line. This establishes the upper bound inequality in (12). So the desired result follows. The proofs for the statements involving \underline{Q}_1 , \bar{Q}_0 , and \underline{Q}_0 are similar by establishing the following bounds:

$$\begin{aligned} \underline{F}_{Y_1}(\underline{Q}_1) - \underline{F}_{Y_1}(\underline{Q}_1-) &\in \left[\frac{p_1}{\bar{c}} \mathbb{P}(Y = \underline{Q}_1 \mid X = 1), \frac{p_1}{\underline{c}} \mathbb{P}(Y = \underline{Q}_1 \mid X = 1) \right] \\ \bar{F}_{Y_0}(\bar{Q}_0) - \bar{F}_{Y_0}(\bar{Q}_0-) &\in \left[\frac{p_0}{1 - \underline{c}} \mathbb{P}(Y = \bar{Q}_0 \mid X = 0), \frac{p_0}{1 - \bar{c}} \mathbb{P}(Y = \bar{Q}_0 \mid X = 0) \right] \\ \underline{F}_{Y_0}(\underline{Q}_0) - \underline{F}_{Y_0}(\underline{Q}_0-) &\in \left[\frac{p_0}{1 - \underline{c}} \mathbb{P}(Y = \underline{Q}_0 \mid X = 0), \frac{p_0}{1 - \bar{c}} \mathbb{P}(Y = \underline{Q}_0 \mid X = 0) \right], \end{aligned} \tag{13}$$

which can be derived by similar steps to those above.

Proof of Part 3: First consider \bar{A}_1 if the denominator is positive, as $\bar{A}_1 = p_1$ is trivially bounded in $[\underline{c}, \bar{c}]$ by Assumption 2 if the denominator becomes zero. By its definition, we have

$$\bar{A}_1 = \frac{p_1 \mathbb{P}(Y = \bar{Q}_1 \mid X = 1)}{\left(\frac{\bar{c} - p_1}{\bar{c}} + F_{Y|X}(\bar{Q}_1 \mid 1) \frac{p_1}{\bar{c}} \right) - \frac{p_1}{\bar{c}} \mathbb{P}(Y < \bar{Q}_1 \mid X = 1)} = \frac{p_1 \mathbb{P}(Y = \bar{Q}_1 \mid X = 1)}{\bar{F}_{Y_1}(\bar{Q}_1) - \bar{F}_{Y_1}(\bar{Q}_1-)}.$$

From the inequality in equation (12), we deduce that $\mathbb{P}(Y = \bar{Q}_1 \mid X = 1) > 0$, and

$$\frac{p_1 \mathbb{P}(Y = \bar{Q}_1 \mid X = 1)}{\bar{F}_{Y_1}(\bar{Q}_1) - \bar{F}_{Y_1}(\bar{Q}_1-)} \in [\underline{c}, \bar{c}].$$

So this concludes that $\bar{A}_1 \in [\underline{c}, \bar{c}]$. Similarly, the results for \underline{A}_1 , $1 - \bar{A}_0$, and $1 - \underline{A}_0$ can be deduced by inequalities (13).

Proof of Part 4: These propensity scores can only take the values $\underline{c}, \bar{c}, \bar{A}_1, \underline{A}_1, 1 - \bar{A}_0$, and $1 - \underline{A}_0$. By Part

3, these values all lie in $[\underline{c}, \bar{c}]$.

Proof of Part 5: First we show that $\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{p_1(Y)} \right] = \underline{F}_{Y_1}(y)$ for all $y \in \mathbb{R}$. The proof for \bar{F}_{Y_1} is similar by interchanging \underline{c} with \bar{c} and thus omitted.

To prove the desired identity, we split the analysis in three cases depending on the value of $y \in \mathbb{R}$. For $y < \underline{Q}_1$, we have $\mathbb{1}(Y \leq y)/p_1(Y) = \mathbb{1}(Y \leq y)/\bar{c}$ and thus

$$\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{p_1(Y)} \right] = \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{\bar{c}} \right] = \frac{F_{Y|X}(y | 1)p_1}{\bar{c}} = \underline{F}_{Y_1}(y).$$

When $y = \underline{Q}_1$, we can write

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{1}(Y \leq \underline{Q}_1)X}{p_1(Y)} \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y < \underline{Q}_1)X}{\bar{c}} \right] + \mathbb{E} \left[\frac{\mathbb{1}(Y = \underline{Q}_1)X}{A_1} \right] \\ &= \frac{\mathbb{P}(Y < \underline{Q}_1 | X = 1)p_1}{\bar{c}} + \mathbb{P}(Y = \underline{Q}_1 | X = 1)p_1 \left(\frac{\mathbb{P}(Y = \underline{Q}_1 | X = 1)p_1}{\underline{F}_{Y_1}(\underline{Q}_1) - \underline{F}_{Y_1}(\underline{Q}_1 -)} \right)^{-1} \\ &= \frac{\mathbb{P}(Y < \underline{Q}_1 | X = 1)p_1}{\bar{c}} + \underline{F}_{Y_1}(\underline{Q}_1) - \underline{F}_{Y_1}(\underline{Q}_1 -) \\ &= \underline{F}_{Y_1}(\underline{Q}_1 -) + \underline{F}_{Y_1}(\underline{Q}_1) - \underline{F}_{Y_1}(\underline{Q}_1 -) \\ &= \underline{F}_{Y_1}(\underline{Q}_1). \end{aligned}$$

Finally, when $y > \underline{Q}_1$, we can write

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{p_1(Y)} \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq \underline{Q}_1)X}{p_1(Y)} \right] + \mathbb{E} \left[\frac{\mathbb{1}(\underline{Q}_1 < Y \leq y)X}{\underline{c}} \right] \\ &= \underline{F}_{Y_1}(\underline{Q}_1) + \frac{\left(F_{Y|X}(y | 1) - F_{Y|X}(\underline{Q}_1 | 1) \right) p_1}{\underline{c}} \\ &= \frac{\underline{c} - p_1}{\underline{c}} + F_{Y|X}(\underline{Q}_1 | 1) \frac{p_1}{\underline{c}} + \frac{\left(F_{Y|X}(y | 1) - F_{Y|X}(\underline{Q}_1 | 1) \right) p_1}{\underline{c}} \\ &= \frac{\underline{c} - p_1}{\underline{c}} + F_{Y|X}(y | 1) \frac{p_1}{\underline{c}} \\ &= \underline{F}_{Y_1}(y). \end{aligned}$$

Thus the desired identity holds for all $y \in \mathbb{R}$.

Next we show the identity $\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - p_0(Y)} \right] = \underline{F}_{Y_0}(y)$ for all $y \in \mathbb{R}$. The proof for \bar{F}_{Y_0} is similar by interchanging \underline{c} with \bar{c} and thus omitted. Similar to above, we split the analysis in three cases depending on the value of $y \in \mathbb{R}$. For $y < \underline{Q}_0$, we have $\mathbb{1}(Y \leq y)/(1 - p_0(Y)) = \mathbb{1}(Y \leq y)/(1 - \underline{c})$ and thus

$$\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - p_0(Y)} \right] = \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - \underline{c}} \right] = \frac{F_{Y|X}(y | 0)p_0}{1 - \underline{c}} = \underline{F}_{Y_0}(y).$$

When $y = \underline{Q}_0$, we can write

$$\begin{aligned}
\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1-X)}{1-p_0(Y)} \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y < \underline{Q}_0)(1-X)}{1-\underline{c}} \right] + \mathbb{E} \left[\frac{\mathbb{1}(Y = \underline{Q}_0)(1-X)}{\underline{A}_0} \right] \\
&= \frac{\mathbb{P}(Y < \underline{Q}_0 | X=0)p_0}{1-\underline{c}} + \mathbb{P}(Y = \underline{Q}_0 | X=0)p_0 \left(\frac{\mathbb{P}(Y = \underline{Q}_0 | X=0)p_0}{\underline{F}_{Y_0}(\underline{Q}_0) - \underline{F}_{Y_0}(\underline{Q}_0^-)} \right)^{-1} \\
&= \frac{\mathbb{P}(Y < \underline{Q}_0 | X=0)p_0}{1-\underline{c}} + \underline{F}_{Y_0}(\underline{Q}_0) - \underline{F}_{Y_0}(\underline{Q}_0^-) \\
&= \underline{F}_{Y_0}(\underline{Q}_0^-) + \underline{F}_{Y_0}(\underline{Q}_0) - \underline{F}_{Y_0}(\underline{Q}_0^-) \\
&= \underline{F}_{Y_0}(\underline{Q}_0).
\end{aligned}$$

When $y > \underline{Q}_0$, we can write

$$\begin{aligned}
\mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1-X)}{1-p_0(Y)} \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq \underline{Q}_0)(1-X)}{1-p_0(Y)} \right] + \mathbb{E} \left[\frac{\mathbb{1}(\underline{Q}_0 < Y \leq y)(1-X)}{1-\bar{c}} \right] \\
&= \underline{F}_{Y_0}(\underline{Q}_0) + \frac{(F_{Y|X}(y|0) - F_{Y|X}(\underline{Q}_0|0))p_0}{1-\bar{c}} \\
&= \frac{p_1 - \bar{c}}{1-\bar{c}} + F_{Y|X}(\underline{Q}_0|0) \frac{p_0}{1-\bar{c}} + \frac{(F_{Y|X}(y|0) - F_{Y|X}(\underline{Q}_0|0))p_0}{1-\bar{c}} \\
&= \frac{p_1 - \bar{c}}{1-\bar{c}} + F_{Y|X}(y|0) \frac{p_0}{1-\bar{c}} \\
&= \underline{F}_{Y_0}(y).
\end{aligned}$$

Thus the desired identity has been established for all $y \in \mathbb{R}$.

Proof of Part 6: We begin by considering the first sequence of equivalences between (a), (b), and (c) for $x = 1$. By Lemma 6.1.

$$\begin{aligned}
\bar{F}_{Y_1}(y_1) \leq \underline{F}_{Y_0}(y_0) &\iff \bar{R}_1(y_1) := \bar{Q}_{Y_0}(\bar{F}_{Y_1}(y_1)) \leq y_0 \\
F_{Y|X}(y_1|1) \leq \underline{F}_{Y_0|X}(y_0|1) &\iff \bar{R}_2(y_1) := \bar{Q}_{Y_0|X}(F_{Y|X}(y_1|1)|1) \leq y_0 \\
\bar{F}_{Y_1|X}(y_1|0) \leq F_{Y_0|X}(y_0|0) &\iff \bar{R}_3(y_1) := Q_{Y|X}(\bar{F}_{Y_1|X}(y_1|0)|0) \leq y_0.
\end{aligned}$$

The equivalence relationship for the statements on the left hand side holds if $\bar{R}_1(y_1) = \bar{R}_2(y_1) = \bar{R}_3(y_1)$ for all $y_1 \in \mathbb{R}$. By direct calculation, we can see that

$$\bar{R}_2(y_1) = \bar{R}_3(y_1) = Q_{Y|X} \left(\min \left\{ \frac{(1-\underline{c})p_1}{p_0\underline{c}} F_{Y|X}(y_1|1), \frac{p_0 - (1-\bar{c})}{p_0\bar{c}} + \frac{p_1(1-\bar{c})}{p_0\bar{c}} F_{Y|X}(y_1|1) \right\} | 0 \right)$$

and thus it remains to show that $\bar{R}_1(y_1) = \bar{R}_2(y_1)$. Recall that

$$\begin{aligned}
\bar{R}_1(y_1) &= \bar{Q}_{Y_0}(\bar{F}_{Y_1}(y_1)) \\
&= Q_{Y|X} \left(\min \left\{ \frac{1-\underline{c}}{p_0} \bar{F}_{Y_1}(y_1), \frac{p_0 - (1-\bar{c})}{p_0} + \frac{1-\bar{c}}{p_0} \bar{F}_{Y_1}(y_1) \right\} | 0 \right).
\end{aligned}$$

We split the proof into two cases. First consider $y_1 < \bar{Q}_1 = Q_{Y|X}(\bar{\tau}_1|1)$. In such case we have

$$F_{Y|X}(y_1|1) < \bar{\tau}_1 = \frac{(\bar{c} - p_1)\underline{c}}{(\bar{c} - \underline{c})p_1} \tag{14}$$

by Lemma 6.1, and

$$\bar{F}_{Y_1}(y_1) = F_{Y|X}(y|1) \frac{p_1}{\underline{c}}. \quad (15)$$

Using equations (14) and (15), it can be verified that

$$\frac{1-\underline{c}}{p_0} \bar{F}_{Y_1}(y_1) < \frac{p_0 - (1-\bar{c})}{p_0} + \frac{1-\bar{c}}{p_0} \bar{F}_{Y_1}(y_1).$$

This implies

$$\bar{R}_1(y_1) = Q_{Y|X} \left(\frac{1-\underline{c}}{p_0} \bar{F}_{Y_1}(y_1) | 0 \right) = Q_{Y|X} \left(\frac{(1-\underline{c})p_1}{p_0 \underline{c}} F_{Y|X}(y_1|1) | 0 \right) \quad \text{if } y_1 < \bar{Q}_1. \quad (16)$$

Next consider $y \geq \bar{Q}_1$. Then we have $F_{Y|X}(y_1|1) \geq \bar{\tau}_1$ by Lemma 6.1, and

$$\bar{F}_{Y_1}(y_1) = \frac{\bar{c} - p_1}{\bar{c}} + F_{Y|X}(y|1) \frac{p_1}{\bar{c}}.$$

These two implications lead to the following inequality

$$\frac{1-\underline{c}}{p_0} \bar{F}_{Y_1}(y_1) \geq \frac{p_0 - (1-\bar{c})}{p_0} + \frac{1-\bar{c}}{p_0} \bar{F}_{Y_1}(y_1),$$

which further implies

$$\begin{aligned} \bar{R}_1(y_1) &= Q_{Y|X} \left(\frac{p_0 - (1-\bar{c})}{p_0} + \frac{1-\bar{c}}{p_0} \bar{F}_{Y_1}(y_1) | 0 \right) \\ &= Q_{Y|X} \left(\frac{p_0 - (1-\bar{c})}{p_0 \bar{c}} + \frac{p_1(1-\bar{c})}{p_0 \bar{c}} F_{Y|X}(y_1|1) | 0 \right) \quad \text{if } y_1 \geq \bar{Q}_1. \end{aligned} \quad (17)$$

By Lemma 6.1, $y_1 < \bar{Q}_1$ is equivalent to $F_{Y|X}(y_1|1) < \bar{\tau}_1$, and it is further equivalent to

$$\frac{(1-\underline{c})p_1}{p_0 \underline{c}} F_{Y|X}(y_1|1) < \frac{p_0 - (1-\bar{c})}{p_0 \bar{c}} + \frac{p_1(1-\bar{c})}{p_0 \bar{c}} F_{Y|X}(y_1|1).$$

From this, (16), and (17), we deduce that

$$\begin{aligned} \bar{R}_2(y_1) &= Q_{Y|X} \left(\min \left\{ \frac{(1-\underline{c})p_1}{p_0 \underline{c}} F_{Y|X}(y_1|1), \frac{p_0 - (1-\bar{c})}{p_0 \bar{c}} + \frac{p_1(1-\bar{c})}{p_0 \bar{c}} F_{Y|X}(y_1|1) \right\} | 0 \right) \\ &= Q_{Y|X} \left(\frac{(1-\underline{c})p_1}{p_0 \underline{c}} F_{Y|X}(y_1|1) | 0 \right) \mathbb{1}(y < \bar{Q}_1) + Q_{Y|X} \left(\frac{\bar{c} - p_1}{p_0 \bar{c}} + \frac{p_1(1-\bar{c})}{p_0 \bar{c}} F_{Y|X}(y_1|1) | 0 \right) \mathbb{1}(y \geq \bar{Q}_1) \\ &= \bar{R}_1(y_1). \end{aligned}$$

Therefore, the desired conclusion for $x = 1$ has been established. Similar arguments can be used to show that the same conclusion also holds for $x = 0$, and to show the second set of equivalences between (d), (e), and (f).

Proof of Part 7: We first consider the equivalence of the statement (a), (b), and (c). We can write

$$\Delta_1 := F_{Y|X}(\bar{Q}_1 - | 1) - F_{Y_0|X}(\underline{Q}_0 - | 1) = F_{Y|X}(\bar{Q}_1 - | 1) - \frac{p_0 \underline{c} F_{Y|X}(\underline{Q}_0 - | 0)}{p_1(1-\underline{c})}$$

From this, we note that

$$\Delta_2 := \bar{F}_{Y_1}(\bar{Q}_1 -) - F_{Y_0}(\underline{Q}_0 -) = \frac{p_1 F_{Y|X}(\bar{Q}_1 - | 1)}{\underline{c}} - \frac{p_0 F_{Y|X}(\underline{Q}_0 - | 0)}{1-\underline{c}} = \frac{p_1}{\underline{c}} \Delta_1,$$

and

$$\Delta_3 := \bar{F}_{Y_1|X}(\bar{Q}_1 - |0) - F_{Y|X}(\underline{Q}_0 - |0) = \frac{p_1(1-\underline{c})}{p_0\underline{c}} F_{Y|X}(\bar{Q}_1 - |1) - F_{Y|X}(\underline{Q}_0 - |0) = \frac{p_1(1-\underline{c})}{p_0\underline{c}} \Delta_1.$$

The desired result follows by noting that Δ_1 , Δ_2 , and Δ_3 all have the same sign. The proof of the equivalence of statements (d), (e), and (f) is similar and thus omitted.

Proof of Part 8: We have that

$$\bar{F}_{Y_1}(\bar{Q}_1 -) = \frac{F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 | 1) - |1)p_1}{\underline{c}} \leq \bar{\tau}_1 \frac{p_1}{\underline{c}} = \frac{\bar{c} - p_1}{\bar{c} - \underline{c}}$$

by Lemma 6.3. Similarly,

$$\bar{F}_{Y_1}(\bar{Q}_1) = \frac{\bar{c} - p_1}{\bar{c}} + \frac{F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 | 1) |1)p_1}{\bar{c}} \geq \frac{\bar{c} - p_1}{\bar{c}} + \bar{\tau}_1 \frac{p_1}{\bar{c}} = \frac{\bar{c} - p_1}{\bar{c} - \underline{c}}.$$

by Lemma 6.2. The other inequalities can be shown in a similar manner. Their derivations are thus omitted.

Proof of Part 9: We have that

$$F_{Y|X}(\bar{Q}_1 - |1) = F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 | 1) - |1) \leq \bar{\tau}_1$$

by Lemma 6.3. Similarly,

$$F_{Y|X}(\bar{Q}_1 |1) = F_{Y|X}(Q_{Y|X}(\bar{\tau}_1 | 1) |1) \geq \bar{\tau}_1$$

by Lemma 6.2. The same arguments can be applied to $\underline{F}_{Y_0|X}(\underline{Q}_0 - |1)$ and $\underline{F}_{Y_0|X}(\underline{Q}_0 |1)$. So we have

$$\underline{F}_{Y_0|X}(\underline{Q}_0 - |1) = F_{Y|X}(Q_{Y|X}(\tau_0 | 0) - |0) \frac{p_0\underline{c}}{p_1(1-\underline{c})} \leq \frac{\tau_0 p_0 \underline{c}}{p_1(1-\underline{c})} = \frac{\underline{c}(\bar{c} - p_1)}{p_1(\bar{c} - \underline{c})} = \bar{\tau}_1.$$

via Lemma 6.3, and

$$\underline{F}_{Y_0|X}(\underline{Q}_0 |1) = \frac{p_1 - \bar{c}}{(1-\bar{c})p_1} + F_{Y|X}(Q_{Y|X}(\tau_0 | 0) |0) \frac{p_0\bar{c}}{p_1(1-\bar{c})} \geq \frac{p_1 - \bar{c}}{(1-\bar{c})p_1} + \frac{\tau_0 p_0 \bar{c}}{p_1(1-\bar{c})} = \bar{\tau}_1.$$

via Lemma 6.2. The proofs for the other inequalities are similar and thus omitted.

Proof of Part 10: We begin by considering the first set of equivalences between (a), (b), and (c) when $x = 1$, and the equivalences for $x = 0$ are identical. By Lemma 6.1, we have the following equivalence relationships:

$$\begin{aligned} \bar{F}_{Y_1}(y_1) + \bar{F}_{Y_0}(y_0) \geq 1 &\iff y_1 \geq \underline{R}_1(y_0) := \underline{Q}_{Y_1}(1 - \bar{F}_{Y_0}(y_0)) \\ F_{Y|X}(y_1 | 1) + \bar{F}_{Y_0|X}(y_0 | 1) \geq 1 &\iff y_1 \geq \underline{R}_2(y_0) := Q_{Y|X}(1 - \bar{F}_{Y_0|X}(y_0 | 1) | 1) \\ \bar{F}_{Y_1|X}(y_1 | 0) + F_{Y|X}(y_0 | 0) \geq 1 &\iff y_1 \geq \underline{R}_3(y_0) := \underline{Q}_{Y_1|X}(1 - F_{Y|X}(y_0 | 0) | 1). \end{aligned}$$

To prove the equivalence of statements on the left, it suffices to show that $\underline{R}_1(y_0) = \underline{R}_2(y_0) = \underline{R}_3(y_0)$ for all

$y_0 \in \mathbb{R}$. First, we directly compute $\underline{R}_2(y_0)$ and $\underline{R}_3(y_0)$:

$$\begin{aligned}
\underline{R}_2(y_0) &= Q_{Y|X}(1 - \overline{F}_{Y_0|X}(y_0 | 1) | 1) \\
&= Q_{Y|X} \left(1 - \min \left\{ \frac{p_0 \bar{c} F_{Y|X}(y_0|0)}{p_1(1-\bar{c})}, \frac{p_1 - \underline{c} + p_0 \underline{c} F_{Y|X}(y_0|0)}{p_1(1-\underline{c})} \right\} | 1 \right) \\
&= Q_{Y|X} \left(\max \left\{ 1 - \frac{p_0 \bar{c} F_{Y|X}(y_0|0)}{p_1(1-\bar{c})}, 1 - \frac{p_1 - \underline{c} + p_0 \underline{c} F_{Y|X}(y_0|0)}{p_1(1-\underline{c})} \right\} | 1 \right) \\
&= Q_{Y|X} \left(\max \left\{ 1 - \frac{p_0 \bar{c} F_{Y|X}(y_0|0)}{p_1(1-\bar{c})}, \frac{p_1(1-\underline{c}) - p_1 + \underline{c} - p_0 \underline{c} F_{Y|X}(y_0|0)}{p_1(1-\underline{c})} \right\} | 1 \right) \\
&= Q_{Y|X} \left(\max \left\{ 1 - \frac{p_0 \bar{c} F_{Y|X}(y_0|0)}{p_1(1-\bar{c})}, \frac{\underline{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\underline{c})} \right\} | 1 \right),
\end{aligned}$$

and

$$\begin{aligned}
\underline{R}_3(y_0) &= Q_{Y_1|X}(1 - F_{Y|X}(y_0|0) | 1) \\
&= Q_{Y|X} \left(\max \left\{ \frac{\underline{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\underline{c})}, \frac{p_1 - \bar{c} + \bar{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\bar{c})} \right\} | 1 \right) \\
&= Q_{Y|X} \left(\max \left\{ \frac{\underline{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\underline{c})}, 1 - \frac{\bar{c} p_0 F_{Y|X}(y_0|0)}{p_1(1-\bar{c})} \right\} | 1 \right).
\end{aligned}$$

Since y_0 was arbitrary, we conclude that $\underline{R}_2(y_0) = \underline{R}_3(y_0)$ for all $y_0 \in \mathbb{R}$.

We next establish that $\underline{R}_1(y_0) = \underline{R}_2(y_0)$ for all $y_0 \in \mathbb{R}$. Note that

$$\begin{aligned}
\underline{R}_1(y_0) &= Q_{Y_1}(1 - \overline{F}_{Y_0}(y_0)) \\
&= Q_{Y|X} \left(\max \left\{ \frac{\underline{c}}{p_1}(1 - \overline{F}_{Y_0}(y_0)), \frac{p_1 - \bar{c}}{p_1} + \frac{\bar{c}}{p_1}(1 - \overline{F}_{Y_0}(y_0)) \right\} | 1 \right) \\
&= Q_{Y|X} \left(\max \left\{ \frac{\underline{c}}{p_1}(1 - \overline{F}_{Y_0}(y_0)), 1 - \frac{\bar{c}}{p_1} \overline{F}_{Y_0}(y_0) \right\} | 1 \right).
\end{aligned}$$

If $y_0 < \overline{Q}_0$, Lemma 6.1 implies

$$\overline{F}_{Y_0}(y_0) = \frac{p_0}{1-\bar{c}} F_{Y|X}(y_0 | 0) \quad \text{and} \quad F_{Y|X}(y_0|0) < \bar{\tau}_0 = \frac{(p_1 - \underline{c})(1 - \bar{c})}{(\bar{c} - \underline{c})p_0}.$$

These two (in)equalities imply that

$$\frac{\underline{c}}{p_1}(1 - \overline{F}_{Y_0}(y_0)) < 1 - \frac{\bar{c}}{p_1} \overline{F}_{Y_0}(y_0).$$

Then it follows that

$$\underline{R}_1(y_0) = Q_{Y|X} \left(1 - \frac{\bar{c}}{p_1} \overline{F}_{Y_0}(y_0) | 1 \right) = Q_{Y|X} \left(1 - \frac{\bar{c} p_0 F_{Y|X}(y_0|0)}{p_1(1-\bar{c})} | 1 \right) \quad \text{if } y_0 < \overline{Q}_0. \quad (18)$$

Similarly, it can be verified that

$$\underline{R}_1(y_0) = Q_{Y|X} \left(\frac{\underline{c}}{p_1}(1 - \overline{F}_{Y_0}(y_0)) | 1 \right) = Q_{Y|X} \left(\frac{\underline{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\underline{c})} | 1 \right) \quad \text{if } y_0 \geq \overline{Q}_0. \quad (19)$$

By Lemma 6.1, $y_0 < \overline{Q}_0$ is equivalent to $F_{Y|X}(y_0|0) < \bar{\tau}_0$, and it is further equivalent to

$$1 - \frac{\bar{c} p_0 F_{Y|X}(y_0|0)}{p_1(1-\bar{c})} > \frac{\underline{c} p_0(1 - F_{Y|X}(y_0|0))}{p_1(1-\underline{c})}.$$

Therefore, we can write

$$\underline{R}_2(y_0) = Q_{Y|X} \left(1 - \frac{\bar{c}p_0 F_{Y|X}(y_0|0)}{p_1(1-\bar{c})} \mid 1 \right) \mathbb{1}(y_0 < \bar{Q}_0) + Q_{Y|X} \left(\frac{cp_0(1 - F_{Y|X}(y_0|0))}{p_1(1-c)} \mid 1 \right) \mathbb{1}(y_0 \geq \bar{Q}_0). \quad (20)$$

By combining (18), (19), and (20), we note that $\underline{R}_2(y_0) = \underline{R}_1(y_0)$ for all $y_0 \in \mathbb{R}$, as desired. The proof for the second set of equivalences between (d), (e), and (f) is similar and thus omitted.

Proof of Part 11: We show the first set of equivalences between (a), (b), and (c), and the proof for the second set of equivalences between (d), (e), and (f) follows similar arguments and thus omitted. First, we expand

$$\Delta'_1 := F_{Y|X}(\bar{Q}_1 \mid 1) + \bar{F}_{Y_0|X}(\bar{Q}_0 - \mid 1) - 1 = F_{Y|X}(\bar{Q}_1 \mid 1) + F_{Y|X}(\bar{Q}_0 - \mid 0) \frac{p_0\bar{c}}{p_1(1-\bar{c})} - 1.$$

Next, note that

$$\begin{aligned} \Delta'_2 &:= \bar{F}_{Y_1}(\bar{Q}_1) + \bar{F}_{Y_0}(\bar{Q}_0 -) - 1 \\ &= \frac{\bar{c} - p_1}{\bar{c}} + F_{Y|X}(\bar{Q}_1 \mid 1) \frac{p_1}{\bar{c}} + F_{Y|X}(\bar{Q}_0 - \mid 0) \frac{p_0}{1-\bar{c}} - 1 \\ &= \frac{p_1}{\bar{c}} \left[F_{Y|X}(\bar{Q}_1 \mid 1) + F_{Y|X}(\bar{Q}_0 - \mid 0) \frac{p_0\bar{c}}{p_1(1-\bar{c})} - 1 \right] \\ &= \frac{p_1}{\bar{c}} \Delta'_1, \end{aligned}$$

and

$$\begin{aligned} \Delta'_3 &:= \bar{F}_{Y_1|X}(\bar{Q}_1|0) + F_{Y|X}(\bar{Q}_0 - \mid 0) - 1 \\ &= \frac{\bar{c} - p_1}{\bar{c}p_0} + F_{Y|X}(\bar{Q}_1 \mid 1) \frac{p_1(1-\bar{c})}{p_0\bar{c}} + F_{Y|X}(\bar{Q}_0 - \mid 0) - 1 \\ &= \frac{p_1(1-\bar{c})}{p_0\bar{c}} \left[F_{Y|X}(\bar{Q}_1 \mid 1) + F_{Y|X}(\bar{Q}_0 - \mid 0) \frac{p_0\bar{c}}{p_1(1-\bar{c})} + \frac{\bar{c} - p_1}{p_1(1-\bar{c})} - \frac{(1-p_1)\bar{c}}{p_1(1-\bar{c})} \right] \\ &= \frac{p_1(1-\bar{c})}{p_0\bar{c}} \left[F_{Y|X}(\bar{Q}_1 \mid 1) + F_{Y|X}(\bar{Q}_0 - \mid 0) \frac{p_0\bar{c}}{p_1(1-\bar{c})} - 1 \right] \\ &= \frac{p_1(1-\bar{c})}{p_0\bar{c}} \Delta'_1. \end{aligned}$$

Therefore, the desired result follows by noting that $\Delta'_1, \Delta'_2, \Delta'_3$ all have the same sign. \square

Proof of Theorem 1. Fix a $w \in \text{supp}(W)$ and $(\varepsilon, \gamma, C_{1,0|1,w}, C_{1,0|0,w}) \in [0, 1]^2 \times \mathcal{C}^2$. We prove this theorem by constructing a probability distribution $\tilde{\mathbb{P}}$ for (Y_1, Y_0, X) conditional on $W = w$ such that for all $y \in \mathbb{R}$, $x \in \{0, 1\}$, and $(y_1, y_0) \in \mathbb{R}^2$, the following conditions hold:

1. $\tilde{\mathbb{P}}(Y_1 \leq y \mid W = w) = \varepsilon F_{Y_1|W}(y \mid w) + (1 - \varepsilon) \bar{F}_{Y_1|W}(y \mid w)$ and $\tilde{\mathbb{P}}(Y_0 \leq y \mid W = w) = \gamma F_{Y_0|W}(y \mid w) + (1 - \gamma) \bar{F}_{Y_0|W}(y \mid w)$;
2. $\tilde{\mathbb{P}}(X = x \mid W = w) = p_{x|w}$;
3. $\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) = F_{Y|X,W}(Y \mid X, w)$;
4. $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid X = x, W = w) = C_{1,0|x,w}(\tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = x, W = w), \tilde{\mathbb{P}}(Y_1 \leq y_0 \mid X = x, W = w))$;
5. $\tilde{\mathbb{P}}(X = 1 \mid Y_x, W = w) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ $\tilde{\mathbb{P}}$ -almost surely.

Condition 1 requires that an arbitrary convex combination of marginal cdf bounds stated in Theorem 1 can be achieved by the constructed measure. Condition 4 then states that any bivariate copula $C_{1,0|x,w}$

is also achievable. Conditions 2 and 3 require the constructed measure generate the same distribution of (Y, X) as the observed data conditional on $W = w$. Finally, Condition 5 requires the marginal c -dependence Assumption 2 to be satisfied for the constructed measure when conditioning on $W = w$. As a result, the constructed measure $\tilde{\mathbb{P}}$ generates the marginal cdfs and copulas in Theorem 1 and satisfies all the requirements in the definition of identified set $\mathcal{I}_0^{\text{marg}}(F_{Y,X,W})$.

For the conciseness of the proof, we write $C_{1,0|x,w}$ as $C_{x,w}$ for $x \in \{0, 1\}$ so that subscripts of copulas denote the conditioning variables.

Let

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) &= xC_{1,w}(F_{Y|X,W}(y_1 | 1, w), F_0(y_0 | 1, w; \gamma))p_{1|w} \\ &\quad + (1-x)C_{0,w}(F_1(y_1 | 0, w; \varepsilon), F_{Y|X,W}(y_0|0, w))p_{0|w}. \end{aligned} \quad (21)$$

where

$$\begin{aligned} F_0(y_0 | 1, w; \gamma) &= \gamma \underline{F}_{Y_0|X,W}(y_0 | 1, w) + (1-\gamma) \overline{F}_{Y_0|X,W}(y_0 | 1, w) \\ F_1(y_1 | 0, w; \varepsilon) &= \varepsilon \underline{F}_{Y_1|X,W}(y_1 | 0, w) + (1-\varepsilon) \overline{F}_{Y_1|X,W}(y_1 | 0, w). \end{aligned}$$

Since convex combinations of cdfs are cdfs, and by Lemma 4.1, both $F_0(\cdot | 1, w; \gamma)$ and $F_1(\cdot | 0, w; \varepsilon)$ are cdfs. By Sklar's Theorem (Nelsen, 2006, Theorem 2.3.3), the expression in (21) is a joint distribution function for (Y_1, Y_0, X) conditional on $W = w$.

For the rest of the proof, we verify conditions 1-5 for the constructed measure $\tilde{\mathbb{P}}$.

Verifying Condition 1: For $y \in \mathbb{R}$, we can see that

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y | W = w) &= \sum_{x \in \{0,1\}} \lim_{y_0 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y, Y_0 \leq y_0, X = x | W = w) \\ &= \lim_{y_0 \rightarrow +\infty} C_{1,w}(F_{Y|X,W}(y_1 | 1, w), F_0(y_0 | 1, w; \gamma))p_{1|w} \\ &\quad + \lim_{y_0 \rightarrow +\infty} C_{0,w}(F_1(y_1 | 0, w; \varepsilon), F_{Y|X,W}(y_0|0, w))p_{0|w} \\ &= C_{1,w}(F_{Y|X,W}(y_1 | 1, w), 1)p_{1|w} + C_{0,w}(F_1(y_1 | 0, w; \varepsilon), 1)p_{0|w} \\ &= F_{Y|X,W}(y_1 | 1, w)p_{1|w} + F_1(y_1 | 0, w; \varepsilon)p_{0|w} \\ &= \varepsilon(F_{Y|X,W}(y_1 | 1, w)p_{1|w} + \underline{F}_{Y_1|X,W}(y_1 | 0, w)p_{0|w}) \\ &\quad + (1-\varepsilon)(F_{Y|X,W}(y_1 | 1, w)p_{1|w} + \overline{F}_{Y_1|X,W}(y_1 | 0, w)p_{0|w}) \\ &= \varepsilon \underline{F}_{Y_1|W}(y | w) + (1-\varepsilon) \overline{F}_{Y_1|W}(y | w). \end{aligned}$$

The third line holds since $C_{x,w}(1, u) = C_{x,w}(u, 1) = u$ for $x \in \{0, 1\}$ and $u \in [0, 1]$. The last line holds by Lemma 4.2.

Likewise,

$$\begin{aligned} \tilde{\mathbb{P}}(Y_0 \leq y | W = w) &= \sum_{x \in \{0,1\}} \lim_{y_1 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y, X = x | W = w) \\ &= \gamma \underline{F}_{Y_0|W}(y | w) + (1-\gamma) \overline{F}_{Y_0|W}(y | w). \end{aligned}$$

Verifying Condition 2: For $x \in \{0, 1\}$, we have that

$$\begin{aligned} \tilde{\mathbb{P}}(X = x | W = w) &= \lim_{y_1, y_0 \rightarrow \infty} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) \\ &= xC_{1,w}(1, 1)p_{1|w} + (1-x)C_{0,w}(1, 1)p_{0|w} \\ &= xp_{1|w} + (1-x)p_{0|w} \\ &= p_{x|w}. \end{aligned}$$

The third equality uses the fact that $C_{x,w}(1, 1) = 1$ for $x \in \{0, 1\}$.

Verifying Condition 3: For $x \in \{0, 1\}$ and $y \in \mathbb{R}$, we have that

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) &= \lim_{y' \rightarrow +\infty} \frac{\tilde{\mathbb{P}}(Y_x \leq y, Y_{1-x} \leq y', X = x \mid W = w)}{\tilde{\mathbb{P}}(X = x \mid W = w)} \\
&= \frac{x C_{1,w}(F_{Y|X,W}(y \mid 1, w), 1) p_{1|w} + (1-x) C_{0,w}(1, F_{Y|X,W}(y \mid 0, w)) p_{0|w}}{p_{x|w}} \\
&= \frac{x F_{Y|X,W}(y \mid 1, w) p_{1|w} + (1-x) F_{Y|X,W}(y \mid 0) p_{0|w}}{p_{x|w}} \\
&= \frac{F_{Y|X,W}(Y \mid X, w) p_{x|w}}{p_{x|w}} \\
&= F_{Y|X,W}(Y \mid X, w).
\end{aligned}$$

The third line holds again by $C_x(1, u) = C_x(u, 1) = u$ for $x \in \{0, 1\}$ and $u \in [0, 1]$. The last line follows by Assumption 1 that $p_{x|w} > 0$ for $x \in \{0, 1\}$.

Verifying Condition 4: First, following similar steps for verifying condition 3, we have

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_x \leq y \mid X = 1-x, W = w) &= \frac{\lim_{y' \rightarrow +\infty} \tilde{\mathbb{P}}(Y_x \leq y, Y_{1-x} \leq y', X = 1-x \mid W = w)}{\tilde{\mathbb{P}}(X = 1-x \mid W = w)} \\
&= \frac{(1-x) C_{1,w}(1, F_0(y_0 \mid 1, w; \gamma)) p_{1|w} + x C_{0,w}(F_1(y_1 \mid 0, w; \epsilon), 1) p_{0|w}}{p_{1-x|w}} \quad (22) \\
&= \frac{(1-x) p_{1|w} F_0(y_0 \mid 1, w; \gamma) + x p_{0|w} F_1(y_1 \mid 0, w; \epsilon)}{p_{1-x|w}} \\
&= (1-x) F_0(y_0 \mid 1, w; \gamma) + x F_1(y_1 \mid 0, w; \epsilon).
\end{aligned}$$

Then for $(y_1, y_0) \in \mathbb{R}^2$, it follows that

$$\begin{aligned}
&\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid X = x) \\
&= x C_{1,w}(F_{Y|X,W}(y_1 \mid 1, w), F_0(y_0 \mid 1, w; \gamma)) + (1-x) C_{0,w}(F_1(y_1 \mid 0, w; \epsilon), F_{Y|X,W}(y_0 \mid 0, w)) \\
&= x C_{1,w}(\tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = 1, W = w), \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = 1, W = w)) \\
&\quad + (1-x) C_{0,w}(\tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = 0, W = w), \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = 0, W = w)) \\
&= C_{x,w}(\tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = x, W = w), \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = x, W = w)).
\end{aligned}$$

The second line holds by Condition 3 and equation (22).

Verifying Condition 5: In this part, we establish an explicit formula of the propensity score function under $\tilde{\mathbb{P}}$ and show that it is contained in $[\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely. To achieve this goal, we divide the analysis into two cases.

Case 1: Consider the case where $p_{1|w} = \underline{c}(w, \eta)$. By direct calculation,

$$\underline{F}_{Y_1|W}(y \mid w) = \bar{F}_{Y_1|W}(y \mid w) = F_{Y|X,W}(y \mid 1, w) \quad \text{and} \quad \underline{F}_{Y_0|W}(y \mid w) = \bar{F}_{Y_0|W}(y \mid w) = F_{Y|X,W}(y \mid 0, w).$$

Based on condition 1 we verified above, we have

$$\tilde{\mathbb{P}}(Y_1 \leq y_1 \mid W = w) = F_{Y|X,W}(y_1 \mid 1, w) \quad \text{and} \quad \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid W = w) = F_{Y|X,W}(y_0 \mid 0, w).$$

Since $p_{1|w} = \underline{c}(w, \eta)$, by Assumption 2, it is straightforwardly to see that $\mathbb{P}(X = 1 \mid Y_1, W = w) = \mathbb{P}(X =$

$1 | Y_0, W = w) = \underline{c}(w, \eta)$ almost surely, which further implies

$$\begin{aligned}\tilde{\mathbb{E}} \left[\frac{\mathbb{1}[Y \leq y_1]X}{\underline{c}(w, \eta)} | W = w \right] &= \mathbb{E} \left[\frac{\mathbb{1}[Y \leq y_1]X}{\underline{c}(w, \eta)} | W = w \right] \\ &= F_{Y|X,W}(y_1 | 1, w) \\ &= \tilde{\mathbb{P}}(Y_1 \leq y_1 | W = w)\end{aligned}$$

and

$$\begin{aligned}\tilde{\mathbb{E}} \left[\frac{\mathbb{1}[Y \leq y_0](1-X)}{1-\underline{c}(w, \eta)} | W = w \right] &= \mathbb{E} \left[\frac{\mathbb{1}[Y \leq y_0](1-X)}{1-\underline{c}(w, \eta)} | W = w \right] \\ &= F_{Y|X,W}(y_0 | 0, w) \\ &= \tilde{\mathbb{P}}(Y_0 \leq y_0 | W = w).\end{aligned}$$

Following Lemma 5, this implies $\tilde{\mathbb{E}}(X | Y_1, W = w) = \tilde{\mathbb{E}}(X | Y_0, W = w) = \underline{c}(w, \eta)$ almost surely under $\tilde{\mathbb{P}}$, which is naturally bounded within $[\underline{c}(w, \eta), \bar{c}(w, \eta)]$, as desired. The proof for the case where $p_{1|w} = \bar{c}(w, \eta)$ follows the same argument by interchanging $\underline{c}(w, \eta)$ with $\bar{c}(w, \eta)$ and thus omitted.

Case 2: Consider the case where $\underline{c}(w, \eta) < p_{1|w} < \bar{c}(w, \eta)$. Define

$$p_1(y, w; \varepsilon) = \frac{1}{\varepsilon \underline{p}_1(y, w)^{-1} + (1-\varepsilon) \bar{p}_1(y, w)^{-1}},$$

where $\bar{p}_1(y, w)$ and $\underline{p}_1(y, w)$ are defined in Appendix A.

By Lemma 7.4, $\underline{p}_1(Y_1, w), \bar{p}_1(Y_1, w) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely. Therefore,

$$p_1(Y_1, w; \varepsilon) = \frac{1}{\varepsilon \underline{p}_1(Y_1, w)^{-1} + (1-\varepsilon) \bar{p}_1(Y_1, w)^{-1}} \leq \frac{1}{\varepsilon \bar{c}(w, \eta)^{-1} + (1-\varepsilon) \bar{c}(w, \eta)^{-1}} = \bar{c}(w, \eta)$$

and

$$p_1(Y_1, w; \varepsilon) = \frac{1}{\varepsilon \underline{p}_1(Y_1, w)^{-1} + (1-\varepsilon) \bar{p}_1(Y_1, w)^{-1}} \geq \frac{1}{\varepsilon \underline{c}(w, \eta)^{-1} + (1-\varepsilon) \underline{c}(w, \eta)^{-1}} = \underline{c}(w, \eta).$$

Therefore $p_1(Y_1, w; \varepsilon) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely.

Next we will show that $\tilde{\mathbb{E}}[X | Y_1, W = w] = p_1(Y_1, w; \varepsilon)$ via Lemma 5 by verifying that

$$\tilde{\mathbb{E}} \left[\frac{\mathbb{1}(Y_1 \leq y)X}{p_1(Y_1, w; \varepsilon)} | W = w \right] = \tilde{\mathbb{P}}(Y_1 \leq y | W = w) = \varepsilon F_{Y_1|W}(y | w) + (1-\varepsilon) \bar{F}_{Y_1|W}(y | w), \text{ for all } y \in \mathbb{R}.$$

To show this, we have the following derivations:

$$\begin{aligned}\tilde{\mathbb{E}} \left[\frac{\mathbb{1}(Y_1 \leq y)X}{p_1(Y_1, w; \varepsilon)} | W = w \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{p_1(Y, w; \varepsilon)} | W = w \right] \\ &= \mathbb{E} \left[\mathbb{1}(Y \leq y)X \left(\frac{\varepsilon}{\underline{p}_1(Y, w)} + \frac{1-\varepsilon}{\bar{p}_1(Y, w)} \right) | W = w \right] \\ &= \varepsilon \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{\underline{p}_1(Y, w)} | W = w \right] + (1-\varepsilon) \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)X}{\bar{p}_1(Y, w)} | W = w \right] \\ &= \varepsilon F_{Y_1|W}(y | w) + (1-\varepsilon) \bar{F}_{Y_1|W}(y | w).\end{aligned}$$

The first equality holds by noting that the distribution of Y_1 conditional $X = 1$ and $W = w$ under $\tilde{\mathbb{P}}$ is the same as the one under the population \mathbb{P} as verified by condition 2. The last equality follows by applying Lemma 7.5.

For the cdf of Y_0 , define

$$p_0(Y_0, w; \gamma) = 1 - \frac{1}{\gamma(1 - \underline{p}_0(Y_0, w))^{-1} + (1 - \gamma)(1 - \bar{p}_0(Y_0, w))^{-1}}.$$

Since $1 - \underline{p}_0(Y_0, w), 1 - \bar{p}_0(Y_0, w) \in [1 - \bar{c}(w, \eta), 1 - \underline{c}(w, \eta)]$ almost surely, we have that

$$\begin{aligned} p_0(Y_0, w; \gamma) &= 1 - \frac{1}{\gamma(1 - \underline{p}_0(Y_0, w))^{-1} + (1 - \gamma)(1 - \bar{p}_0(Y_0, w))^{-1}} \\ &\leq 1 - \frac{1}{\gamma(1 - \bar{c}(w, \eta))^{-1} + (1 - \gamma)(1 - \bar{c}(w, \eta))^{-1}} \\ &= \bar{c}(w, \eta), \end{aligned}$$

and

$$\begin{aligned} p_0(Y_0, w; \gamma) &= 1 - \frac{1}{\gamma(1 - \underline{p}_0(Y_0, w))^{-1} + (1 - \gamma)(1 - \bar{p}_0(Y_0, w))^{-1}} \\ &\geq 1 - \frac{1}{\gamma(1 - \underline{c}(w, \eta))^{-1} + (1 - \gamma)(1 - \underline{c}(w, \eta))^{-1}} \\ &= \underline{c}(w, \eta). \end{aligned}$$

Therefore, $p_0(Y_0, w; \eta) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely. We can also see that

$$\begin{aligned} \tilde{\mathbb{E}} \left[\frac{\mathbb{1}(Y_0 \leq y)(1 - X)}{1 - p_0(Y_0, w; \gamma)} \mid W = w \right] &= \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - p_0(Y, w; \eta)} \mid W = w \right] \\ &= \mathbb{E} \left[\mathbb{1}(Y \leq y)(1 - X) \left(\frac{\gamma}{1 - \underline{p}_0(Y_0, w)} + \frac{1 - \gamma}{1 - \bar{p}_0(Y_0, w)} \right) \mid W = w \right] \\ &= \gamma \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - \underline{p}_0(Y_0, w)} \mid W = w \right] + (1 - \gamma) \mathbb{E} \left[\frac{\mathbb{1}(Y \leq y)(1 - X)}{1 - \bar{p}_0(Y_0, w)} \mid W = w \right] \\ &= \gamma F_{Y_0|W}(y | w) + (1 - \gamma) \bar{F}_{Y_0|W}(y | w), \end{aligned}$$

where the last equality follows by Lemma 7.5. Therefore, by Lemma 5, $\tilde{\mathbb{P}}(X = 1 \mid Y_0, W = w) = p_0(Y_0, w; \eta) \in [\underline{c}(w, \eta), \bar{c}(w, \eta)]$ almost surely, which concludes the proof. \square

C.2 Proof of Theorem 2

This appendix provides a proof of Theorem 2 and all of its auxiliary lemmas. First, we define four latent propensity score functions. For $w \in \text{supp}(W)$, let

$$p^{ul}(y_1, y_0, w; B) = \begin{cases} \underline{c} & \text{if } y_1 \leq \bar{Q}_1, y_0 \leq \underline{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \underline{Q}_0) \\ B & \text{if } (y_1, y_0) = (\bar{Q}_1, \underline{Q}_0) \\ \bar{c} & \text{if } y_1 \geq \bar{Q}_1, y_0 \geq \underline{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \underline{Q}_0) \\ p_{1|w} & \text{otherwise,} \end{cases} \quad (23)$$

$$p^{uu}(y_1, y_0, w; B) = \begin{cases} \underline{c} & \text{if } y_1 \leq \bar{Q}_1, y_0 \geq \bar{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \bar{Q}_0) \\ B & \text{if } (y_1, y_0) = (\bar{Q}_1, \bar{Q}_0) \\ \bar{c} & \text{if } y_1 \geq \bar{Q}_1, y_0 \leq \bar{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \bar{Q}_0) \\ p_{1|w} & \text{otherwise,} \end{cases} \quad (24)$$

$$p^{lu}(y_1, y_0, w; B) = \begin{cases} \bar{c} & \text{if } y_1 \leq \bar{Q}_1, y_0 \leq \underline{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \underline{Q}_0) \\ B & \text{if } (y_1, y_0) = (\bar{Q}_1, \underline{Q}_0) \\ \underline{c} & \text{if } y_1 \geq \bar{Q}_1, y_0 \geq \underline{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \underline{Q}_0) \\ p_{1|w} & \text{otherwise,} \end{cases} \quad (25)$$

$$p^{lu}(y_1, y_0, w; B) = \begin{cases} \bar{c} & \text{if } y_1 \leq \bar{Q}_1, y_0 \geq \bar{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \bar{Q}_0) \\ B & \text{if } (y_1, y_0) = (\bar{Q}_1, \bar{Q}_0) \\ \underline{c} & \text{if } y_1 \geq \bar{Q}_1, y_0 \leq \bar{Q}_0, (y_1, y_0) \neq (\bar{Q}_1, \bar{Q}_0) \\ p_{1|w} & \text{otherwise.} \end{cases} \quad (26)$$

By appropriately specifying the constant B in these propensity scores, we can show that they correspond to the propensity scores $\mathbb{P}(X = 1 \mid Y_1, Y_0, W = w)$ under joint c -dependence for all four pairs of cdf bounds. Before showing this, we state and prove three auxiliary lemmas.

Lemma 8. Let $w \in \text{supp}(W)$. Suppose $m(\cdot)$ is a Borel measurable function and $\mathbb{P}(m(Y_1, Y_0) > \delta \mid W = w) = 1$ for some $\delta > 0$. The following statements are equivalent:

1. Conditional on $W = w$, the following statement holds almost surely:

$$m(Y_1, Y_0) = \mathbb{E}[X \mid Y_1, Y_0, W = w]. \quad (27)$$

2. For all $(y_1, y_0) \in \mathbb{R}^2$, the following equality holds:

$$\mathbb{E}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)m(Y_1, Y_0) \mid W = w] = \mathbb{P}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w). \quad (28)$$

Proof of Lemma 8. We first show (27) implies (28), note that

$$\begin{aligned} \mathbb{P}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w) &= \mathbb{E}[\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]X \mid W = w] \\ &= \mathbb{E}(\mathbb{E}[\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]X \mid Y_1, Y_0, W = w] \mid W = w) \\ &= \mathbb{E}(\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]\mathbb{E}[X \mid Y_1, Y_0, W = w] \mid W = w) \\ &= \mathbb{E}(\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]m(Y_1, Y_0) \mid W = w). \end{aligned} \quad (29)$$

where we use the law of iterated expectation in the second line and use (27) in the last line of derivation.

Next, we show that (28) implies (27). To establish this result, we first note a few key facts:

1. Following from the last two lines of (29), the law of iterated expectations implies

$$\mathbb{E}[\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]\mathbb{E}(X \mid Y_1, Y_0, W = w) \mid W = w] = \mathbb{E}[\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]m(Y_1, Y_0) \mid W = w]$$

for each $(y_1, y_0) \in \mathbb{R}^2$.

2. For $(y_1, y_0) \in \mathbb{R}^2$, define the preimage from a half-space on \mathbb{R}^2 :

$$I_{y_1, y_0} = \{\omega \in \Omega : Y_1(\omega) \leq y_1, Y_0(\omega) \leq y_0\}$$

and let $\mathcal{A}_2 := \{I_{y_1, y_0} : (y_1, y_0) \in \mathbb{R}^2\}$. Similar to the proof of lemma 5, the class of sets \mathcal{A}_2 is a π -system.

3. The sample space can be written as a countable union of sets in \mathcal{A}_2 :

$$\Omega = \{\omega \in \Omega : Y_1(\omega) < \infty, Y_0(\omega) < \infty\} = \bigcup_{n=1}^{\infty} I_{n, n}.$$

4. The random variable $m(Y_1, Y_0)$ is measurable with respect to the σ -algebra generated by (Y_1, Y_0) due to the Borel measurability of $m(\cdot)$, and it is integrable since $\mathbb{E}(m(Y_1, Y_0) | W = w) = \mathbb{P}(X = 1 | W = w) < \infty$ by sending y_1 and y_0 to infinity in (29).
5. The σ -algebra generated by \mathcal{A}_2 equals the σ -algebra generated by (Y_1, Y_0) , i.e.,

$$\sigma(\mathcal{A}_2) = \sigma(Y_1, Y_0).$$

To show this, define the mapping $f : \Omega \rightarrow \mathbb{R}^2$ as $f(\omega) \mapsto (Y_1(\omega), Y_0(\omega))$ and $\mathcal{F} = \{(-\infty, y_1] \times (-\infty, y_0] : (y_1, y_0) \in \mathbb{R}^2\}$. Note that

$$\sigma(\mathcal{A}_2) = \sigma(f^{-1}(\mathcal{F})) = f^{-1}(\sigma(\mathcal{F})).$$

Since the Borel σ -algebra on \mathbb{R}^2 can be generated by elements in \mathcal{F} , we have $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R}^2)$. This implies

$$f^{-1}(\sigma(\mathcal{F})) = f^{-1}(\mathcal{B}(\mathbb{R}^2)) := \sigma(Y_1, Y_0).$$

Therefore the desired conclusion holds.

Given the above results, it follows by Billingsley (1995, Theorem 34.1) that

$$\mathbb{E}[X | Y_1, Y_0, W = w] = m(Y_1, Y_0),$$

almost surely conditional on $W = w$, as desired. □

Lemma 9. Let $w \in \text{supp}(W)$. Consider a probability distribution defined by

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) &= \min\{F_{Y|X,W}(y_1 | 1, w), \underline{F}_{Y_0|X,W}(y_0 | 1, w)\} p_{1|w} x \\ &\quad + \min\{\bar{F}_{Y_1|X,W}(y_1 | 0, w), F_{Y|X,W}(y_0|0, w)\} p_{0|w} (1 - x), \end{aligned}$$

then

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) = \min\{\bar{F}_{Y_1|W}(y_1 | w), \underline{F}_{Y_0|W}(y_0 | w)\}.$$

Also for the following distribution,

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) &= \min\{F_{Y|X,W}(y_1 | 1, w), \bar{F}_{Y_0|X,W}(y_0 | 1, w)\} p_{1|w} x \\ &\quad + \min\{\underline{F}_{Y_1|X,W}(y_1 | 0, w), F_{Y|X,W}(y_0|0, w)\} p_{0|w} (1 - x) \end{aligned}$$

implies

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) = \min\{\underline{F}_{Y_1}(y_1 | w), \bar{F}_{Y_0}(y_0 | w)\}.$$

Proof of Lemma 9. Consider the first statement with $p_{1|w} = \underline{c}$. Then it follows that $\bar{F}_{Y_1|X,W}(y_1 | 0, w) = \bar{F}_{Y_1|W}(y_1 | w) = F_{Y|X,W}(y_1 | 1, w)$ and $\underline{F}_{Y_0|X,W}(y_0 | 1, w) = \underline{F}_{Y_0|W}(y_0 | w) = F_{Y|X,W}(y_0|0, w)$. Therefore,

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) = \min\{F_{Y|X,W}(y_1 | 1, w), F_{Y|X,W}(y_0|0, w)\} p_{x|w}.$$

This implies

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) &= \min\{F_{Y|X,W}(y_1 | 1, w), F_{Y|X,W}(y_0|0, w)\} \\ &= \min\{\bar{F}_{Y_1|W}(y_1 | w), \underline{F}_{Y_0|W}(y_0 | w)\} \end{aligned}$$

as desired. The proof for the case where $p_{1|w} = \bar{c}$ follows the same arguments and thus omitted.

Next consider $\underline{c} < p_{1|w} < \bar{c}$. We have that

$$\begin{aligned}
& \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) \\
&= \min\{F_{Y|X,W}(y_1 \mid 1, w), \underline{F}_{Y_0|X,W}(y_0 \mid 1, w)\}p_{1|w} + \min\{\bar{F}_{Y_1|X,W}(y_1 \mid 0, w), F_{Y|X,W}(y_0|0, w)\}p_{0|w} \\
&= F_{Y|X,W}(y_1 \mid 1, w)p_{1|w}\mathbb{1}(\underline{F}_{Y_0|X,W}(y_0 \mid 1, w) \geq F_{Y|X,W}(y_1 \mid 1, w)) \\
&\quad + \underline{F}_{Y_0|X,W}(y_0 \mid 1, w)p_{1|w}\mathbb{1}(\underline{F}_{Y_0|X,W}(y_0 \mid 1, w) < F_{Y|X,W}(y_1 \mid 1, w)) \\
&\quad + \bar{F}_{Y_1|X,W}(y_1 \mid 0, w)p_{0|w}\mathbb{1}(F_{Y|X,W}(y_0|0, w) \geq \bar{F}_{Y_1|X,W}(y_1 \mid 0, w)) \\
&\quad + F_{Y|X,W}(y_0|0, w)p_{0|w}\mathbb{1}(F_{Y|X,W}(y_0|0, w) < \bar{F}_{Y_1|X,W}(y_1 \mid 0, w)) \\
&= (F_{Y|X,W}(y_1 \mid 1, w)p_{1|w} + \bar{F}_{Y_1|X,W}(y_1 \mid 0, w)p_{0|w})\mathbb{1}[\underline{F}_{Y_0|W}(y_0 \mid w) \geq \bar{F}_{Y_1|W}(y_1 \mid w)] \\
&\quad + \left(\underline{F}_{Y_0|X,W}(y_0 \mid 1, w)p_{1|w} + F_{Y|X,W}(y_0|0, w)p_{0|w}\right)\mathbb{1}[\underline{F}_{Y_0|W}(y_0 \mid w) < \bar{F}_{Y_1|W}(y_1 \mid w)] \\
&= \min\{\bar{F}_{Y_1|W}(y_1 \mid w), \underline{F}_{Y_0|W}(y_0 \mid w)\}.
\end{aligned}$$

The third equality follows by the first set of equivalences in Lemma 7.6 after setting $x = 1$. The last equality follows by Lemma 4.2. The second statement follows similar arguments but instead uses Lemma 7.6 by setting $x = 0$. Therefore, the proof is complete. \square

Lemma 10. Let $w \in \text{supp}(W)$. Consider a probability distribution defined by

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x \mid W = w) &= \max\{F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0\}p_{1|w}x \\
&\quad + \max\{\bar{F}_{Y_1|X,W}(y_1 \mid 0, w) + F_{Y|X,W}(y_0|0, w) - 1, 0\}p_{0|w}(1 - x)
\end{aligned}$$

then

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) = \max\{\bar{F}_{Y_1|W}(y_1 \mid w) + \bar{F}_{Y_0|W}(y_0 \mid w) - 1, 0\}.$$

Also for the following distribution,

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x \mid W = w) &= \max\{F_{Y|X,W}(y_1 \mid 1, w) + \underline{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0\}p_{1|w}x \\
&\quad + \max\{\underline{F}_{Y_1|X,W}(y_1 \mid 0, w) + F_{Y|X,W}(y_0|0, w) - 1, 0\}p_{0|w}(1 - x)
\end{aligned}$$

implies

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) = \max\{\underline{F}_{Y_1|W}(y_1 \mid w) + \underline{F}_{Y_0|W}(y_0 \mid w) - 1, 0\}.$$

Proof of Lemma 10. Consider the first statement. Similar arguments from the proof of Lemma 9 can be used to establish the desired result for $p_{1|w} = \underline{c}$ or $p_{1|w} = \bar{c}$. Thus we consider the case where $\underline{c} < p_{1|w} < \bar{c}$. Then we have that

$$\begin{aligned}
& \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) \\
&= \max\{F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0\}p_{1|w} \\
&\quad + \max\{\bar{F}_{Y_1|X,W}(y_1 \mid 0, w) + F_{Y|X,W}(y_0|0, w) - 1, 0\}p_{0|w} \\
&= \max\{p_{1|w}(F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1), 0\} \\
&\quad + \max\{p_{0|w}(\bar{F}_{Y_1|X,W}(y_1 \mid 0, w) + F_{Y|X,W}(y_0|0, w) - 1), 0\} \\
&= \max\left\{\sum_{x=0,1} p_x F_{Y|X,W}(y_1 \mid x, w) + p_{0|w}\bar{F}_{Y_1|X,W}(y_1 \mid 0, w) + p_{1|w}\bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - (p_{1|w} + p_{0|w}), 0\right\} \\
&= \max\{\bar{F}_{Y_1|W}(y_1 \mid w), \bar{F}_{Y_0|W}(y_0 \mid w) - 1, 0\}.
\end{aligned}$$

The second equality follows by the first set of equivalences in Lemma 7.10 by setting $x = 1$. The last equality holds by Lemma 4.2. The second statement follows similar arguments but instead uses Lemma 7.10 on the

second set of equivalences regarding lower bounds of cdfs. Therefore, the proof is complete. \square

Lemma 11. Let assumptions 1 and 3 hold. Let $\overline{\mathcal{C}}_{1,0|X,W}$ and $\underline{\mathcal{C}}_{1,0|X,W}$ denote classes of comonotonic (and counter-monotonic) copulas where $\overline{\mathcal{C}}_{1,0|x,w}(u, v) = \min\{u, v\}$ and $\underline{\mathcal{C}}_{1,0|x,w}(u, v) = \max\{u + v - 1, 0\}$ for all $(x, w) \in \{0, 1\} \times \text{supp}(W)$. Then each of the following terms is contained in the identified set $\mathcal{I}_0^j(F_{Y,X,W})$:

1. $(\overline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, \overline{\mathcal{C}}_{1,0|X,W})$;
2. $(\overline{F}_{Y_1|W}, \overline{F}_{Y_0|W}, \underline{\mathcal{C}}_{1,0|X,W})$;
3. $(\underline{F}_{Y_1|W}, \overline{F}_{Y_0|W}, \overline{\mathcal{C}}_{1,0|X,W})$;
4. $(\underline{F}_{Y_1|W}, \underline{F}_{Y_0|W}, \underline{\mathcal{C}}_{1,0|X,W})$.

Proof of Lemma 11. Proof of Part 1: Fix a $w \in \text{supp}(W)$. We prove the first statement by constructing a probability distribution $\tilde{\mathbb{P}}$ for (Y_1, Y_0, X) conditional on $W = w$ such that for all $y \in \mathbb{R}$ and $x \in \{0, 1\}$, we have

1. $\tilde{\mathbb{P}}(Y_1 \leq y | W = w) = \overline{F}_{Y_1|W}(y | w)$ and $\tilde{\mathbb{P}}(Y_0 \leq y | W = w) = \underline{F}_{Y_0|W}(y | w)$;
2. $\tilde{\mathbb{P}}(X = x | W = w) = p_{x|w}$;
3. $\tilde{\mathbb{P}}(Y_x \leq y | X = x, W = w) = F_{Y|x,W}(Y | X, w)$;
4. $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | X = x, W = w) = \min \left\{ \tilde{\mathbb{P}}(Y_1 \leq y_1 | X = x, W = w), \tilde{\mathbb{P}}(Y_0 \leq y_0 | X = x, W = w) \right\}$;
5. $\tilde{\mathbb{P}}(X = 1 | Y_1, Y_0, W = w) \in [\underline{c}, \overline{c}]$ for $\tilde{\mathbb{P}}$ -almost surely.

Similar to the arguments in the proof of Theorem 1, Conditions 1–5 ensures that the constructed distribution $\tilde{\mathbb{P}}$ generates the desired marginal cdfs and copulas in Lemma 11.1 and satisfies all the requirements in the definition of identified set $\mathcal{I}_0^j(F_{Y,X,W})$.

Let

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) &= x \min\{F_{Y|x,W}(y_1 | 1, w), \underline{F}_{Y_0|x,W}(y_0 | 1, w)\} p_{1|w} \\ &\quad + (1 - x) \min\{\overline{F}_{Y_1|x,W}(y_1 | 0, w), F_{Y|x,W}(y_0 | 0, w)\} p_{0|w}. \end{aligned} \quad (30)$$

By Lemma 4.1, $\underline{F}_{Y_0|x,W}(y_0 | 1, w)$ and $\overline{F}_{Y_1|x,W}(y_1 | 0, w)$ are cdfs. Also note that $(u, v) \mapsto \min\{u, v\}$ is the comonotonic copula. By Sklar's Theorem, $\tilde{\mathbb{P}}$ is a joint distribution function for (Y_1, Y_0, X) conditional on $W = w$.

Following the same steps as in the proof of Theorem 1, it can be shown that conditions 1–4 are satisfied because the distribution in (30) is the same as in (21) but for a specific rather than an arbitrary choice of copulas. By Lemma 9, $\tilde{\mathbb{P}}$ implies the following co-monotonic joint distribution of (Y_1, Y_0) :

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) = \min \left\{ \overline{F}_{Y_1|W}(y_1 | w), \underline{F}_{Y_0|W}(y_0 | w) \right\}. \quad (31)$$

To show condition 5 holds, and thus complete the proof, we construct a function p^{ul} such that $p^{ul}(Y_1, Y_0) = \tilde{\mathbb{E}}[X | Y_1, Y_0, W = w]$ and $p^{ul}(Y_1, Y_0, w) \in [\underline{c}, \overline{c}]$ almost surely under $\tilde{\mathbb{P}}$.

First consider $p_{1|w} = \underline{c}$, then we have

$$\underline{F}_{Y_0|W}(y_0 | w) = \underline{F}_{Y_0|x,W}(y_0 | 1, w) = F_{Y|x,W}(y_0 | 0, w)$$

and

$$\overline{F}_{Y_1|W}(y_1 | w) = \overline{F}_{Y_1|x,W}(y_1 | 0, w) = F_{Y|x,W}(y_1 | 1, w).$$

This implies the following derivation

$$\begin{aligned} \tilde{\mathbb{E}}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) \underline{c} | W = w] &= p_{1|w} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) \\ &= p_{1|w} \min \left\{ F_{Y|x,W}(y_1 | 1, w), F_{Y|x,W}(y_0 | 0, w) \right\} \\ &= \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 | W = w). \end{aligned}$$

The first line holds by $\underline{c} = p_{1|w}$, the second line holds by (31), and the last line holds by (30). Following Lemma 8, we conclude that $\tilde{\mathbb{P}}(X = 1 | Y_1, Y_0, W = w) = \underline{c}$ almost surely, which is naturally bounded between \underline{c} and \bar{c} . The proof of the case where $p_{1|w} = \bar{c}$ is similar and thus omitted.

Next consider $\underline{c} < p_{1|w} < \bar{c}$. Let $p^{ul}(Y_1, Y_0, w) = p^{ul}(Y_1, Y_0, w; B^{ul})$ defined in (23), where

$$B^{ul} = \frac{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0, X = 1 | W = w)}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | W = w)}$$

whenever $\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | W = w) > 0$. Let $B^{ul} = p_{1|w}$ otherwise.

We verify that $B^{ul} \in [\underline{c}, \bar{c}]$ if the denominator is nonzero.

First, note that the denominator of B^{ul} can be expanded below

$$\begin{aligned} & \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | W = w) \\ &= \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \underline{Q}_0 | W = w) - \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 < \underline{Q}_0 | W = w) \\ & \quad - \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 \leq \underline{Q}_0 | W = w) + \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 < \underline{Q}_0 | W = w) \\ &= \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w)\} - \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 - |w)\} \\ & \quad - \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w)\} + \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \underline{F}_{Y_0|W}(\underline{Q}_0 - |w)\}, \end{aligned}$$

where the second equality holds via (31). By Lemma 7.8, this expression simplifies to

$$\begin{aligned} & \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | W = w) \\ &= \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w)\} - \underline{F}_{Y_0|W}(\underline{Q}_0 - |w) \\ & \quad - \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \underline{F}_{Y_0|W}(\underline{Q}_0 - |w)\} \\ &= \min\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w)\} - \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \underline{F}_{Y_0|W}(\underline{Q}_0 - |w)\}. \end{aligned}$$

Second, we expand the numerator of B^{ul} . We have that

$$\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0, X = 1 | W = w) = \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | X = 1, W = w)p_{1|w}$$

and that

$$\begin{aligned} & \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 | X = 1, W = w) \\ &= \min\{F_{Y|X,W}(\bar{Q}_1 | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w)\} - \min\{F_{Y|X,W}(\bar{Q}_1 | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w)\} \\ & \quad - \min\{F_{Y|X,W}(\bar{Q}_1 - | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w)\} + \min\{F_{Y|X,W}(\bar{Q}_1 - | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w)\} \\ &= \min\{F_{Y|X,W}(\bar{Q}_1 | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w)\} - \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w) \\ & \quad - F_{Y|X,W}(\bar{Q}_1 - | 1, w) + \min\{F_{Y|X,W}(\bar{Q}_1 - | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w)\} \\ &= \min\{F_{Y|X,W}(\bar{Q}_1 | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w)\} - \max\{F_{Y|X,W}(\bar{Q}_1 - | 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 - | 1, w)\}, \end{aligned}$$

where the second to last equality follows from Lemma 7.9.

From Part 6 and 7 of Lemma 7, we observe that B^{ul} can take four possible values as follows:

$$\begin{aligned} B^{ul} &= \frac{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 \mid X = 1, W = w)p_{1|w}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 \mid W = w)} \\ &= \frac{(F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 - \mid 1, w))p_{1|w}}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w), \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w) > \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w) \end{array} \right) \end{aligned} \quad (32)$$

$$+ \frac{(F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \underline{F}_{Y_0|X,W}(\underline{Q}_0 - \mid 1, w))p_{1|w}}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w), \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w) \end{array} \right) \quad (33)$$

$$+ \frac{(\underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 - \mid 1, w))p_{1|w}}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) > \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w), \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w) > \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w) \end{array} \right) \quad (34)$$

$$+ \frac{(\underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) - \underline{F}_{Y_0|X,W}(\underline{Q}_0 - \mid 1, w))p_{1|w}}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) > \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w), \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w) \end{array} \right). \quad (35)$$

All the terms have positive denominators since we focus on the case where $\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \underline{Q}_0 \mid W = w) > 0$. As shown in Lemma 7.4, terms (32) and (35) lie in $[\underline{c}, \bar{c}]$.

Next we examine the term (33), which can be written as follows

$$\begin{aligned} &\frac{(F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \underline{F}_{Y_0|X,W}(\underline{Q}_0 - \mid 1, w))p_{1|w}}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} \\ &= \frac{F_{Y|X,W}(\bar{Q}_1 \mid 1, w)p_{1|w} - F_{Y|X,W}(\underline{Q}_0 - \mid 0, w) \frac{p_{0|w}\underline{c}}{1-\underline{c}}}{\frac{\bar{c}-p_{1|w}}{\bar{c}} + F_{Y|X,W}(\bar{Q}_1 \mid 1, w) \frac{p_{1|w}}{\bar{c}} - F_{Y|X,W}(\underline{Q}_0 - \mid 0, w) \frac{p_{0|w}}{1-\underline{c}}} \\ &= \underline{c} + \frac{(\bar{c} - \underline{c})p_{1|w}}{\bar{c}} \frac{F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1}{\frac{\bar{c}-p_{1|w}}{\bar{c}} + F_{Y|X,W}(\bar{Q}_1 \mid 1, w) \frac{p_{1|w}}{\bar{c}} - F_{Y|X,W}(\underline{Q}_0 - \mid 0, w) \frac{p_{0|w}}{1-\underline{c}}} \\ &= \underline{c} + \frac{(\bar{c} - \underline{c})p_{1|w}}{\bar{c}} \frac{F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} \\ &\geq \underline{c} \end{aligned}$$

where the last line follows by $\bar{c} \geq \underline{c}$ and $F_{Y|X,W}(\bar{Q}_1 \mid 1, w) \geq \bar{\tau}_1$ via Lemma 6.2. Also note that

$$\frac{(F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \underline{F}_{Y_0|X,W}(\underline{Q}_0 - \mid 1, w))p_{1|w}}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} = \bar{c} + \frac{p_{0|w}(\bar{c} - \underline{c})}{1 - \underline{c}} \frac{F_{Y|X,W}(\underline{Q}_0 - \mid 0, w) - \underline{\tau}_0}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - \mid w)} \leq \bar{c},$$

where the inequality follows by $\bar{c} \geq \underline{c}$ and $F_{Y|X,W}(\underline{Q}_0 - \mid 0, w) \leq \underline{\tau}_0$ via Lemma 6.3. Thus we have shown the term (33) is bounded within $[\underline{c}, \bar{c}]$.

Then consider the term (34). Following the same arguments, we have

$$\frac{(\underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 - \mid 1, w))p_{1|w}}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} = \underline{c} + \frac{p_{0|w}(\bar{c} - \underline{c})}{1 - \bar{c}} \frac{F_{Y|X,W}(\underline{Q}_0 \mid 0, w) - \underline{\tau}_0}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} \geq \underline{c},$$

and

$$\frac{(\underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 - \mid 1, w))p_{1|w}}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} = \bar{c} + \frac{p_{1|w}(\bar{c} - \underline{c})}{\underline{c}} \frac{F_{Y|X,W}(\bar{Q}_1 - \mid 1, w) - \bar{\tau}_1}{\underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - \mid w)} \leq \bar{c}.$$

So we have shown that all the four terms (32)–(35) are bounded within $[\underline{c}, \bar{c}]$, thus concluding $B^{ul} \in [\underline{c}, \bar{c}]$, which then establishes $p^{ul}(Y_1, Y_0, w) \in [\underline{c}, \bar{c}]$ almost surely.

To finish this proof, we demonstrate that $\tilde{\mathbb{E}}[X \mid Y_1, Y_0, W = w] = p^{ul}(Y_1, Y_0, w)$ almost surely. To do so,

we use Lemma 8 and show that

$$\begin{aligned}\tilde{\mathbb{E}}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)p^{ul}(Y_1, Y_0, w) | W = w] &= \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 | W = w) \\ &= p_{1|w} \min\{F_{Y|X,W}(y_1 | 1, w), \underline{F}_{Y_0|X,W}(y_0 | 1, w)\}\end{aligned}\quad (36)$$

for all $(y_1, y_0) \in \mathbb{R}^2$. To complete the proof, we break this up into different cases.

(Part 1) Case 1: $y_1 < \bar{Q}_1$ and $y_0 < \underline{Q}_0$.

In this case, $p^{ul}(Y_1, Y_0, w)\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0] = \underline{c}\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0]$. Thus we have

$$\begin{aligned}\tilde{\mathbb{E}}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)p^{ul}(Y_1, Y_0, w) | W = w] &= \underline{c} \min\left\{\bar{F}_{Y_1|W}(y_1 | w), \underline{F}_{Y_0|W}(y_0 | w)\right\} \\ &= \underline{c} \min\left\{F_{Y|X,W}(y_1 | 1, w)\frac{p_{1|w}}{\underline{c}}, F_{Y|X,W}(y_0|0, w)\frac{p_{0|w}}{1-\underline{c}}\right\} \\ &= \min\left\{F_{Y|X,W}(y_1 | 1, w)p_{1|w}, F_{Y|X,W}(y_0|0, w)\frac{p_{0|w}\underline{c}}{1-\underline{c}}\right\} \\ &= p_{1|w} \min\left\{F_{Y|X,W}(y_1 | 1, w), \underline{F}_{Y_0|X,W}(y_0 | 1, w)\right\}.\end{aligned}$$

The second line holds by the assumption that $y_1 < \bar{Q}_1$ and $y_0 < \underline{Q}_0$. Therefore, we have shown that (36) holds.

(Part 1) Case 2: $y_1 \geq \bar{Q}_1$ and $y_0 < \underline{Q}_0$.

First, note that the joint cdf from (31) implies

$$\tilde{\mathbb{P}}(Y_1 > \bar{Q}_1, Y_0 < \underline{Q}_0 | W = w) = \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 > \underline{Q}_0 | W = w) = 0. \quad (37)$$

These equalities follow by

$$\begin{aligned}\tilde{\mathbb{P}}(Y_1 > \bar{Q}_1, Y_0 < \underline{Q}_0 | W = w) &= \tilde{\mathbb{P}}(Y_0 < \underline{Q}_0 | W = w) - \tilde{\mathbb{P}}(Y_0 < \underline{Q}_0, Y_1 \leq \bar{Q}_1 | W = w) \\ &= \underline{F}_{Y_0|W}(\underline{Q}_0 - |w) - \min\left\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 - |w)\right\} \\ &= \underline{F}_{Y_0|W}(\underline{Q}_0 - |w) - \underline{F}_{Y_0|W}(\underline{Q}_0 - |w) \\ &= 0,\end{aligned}$$

where the third line holds by Lemma 7.8. Similarly,

$$\begin{aligned}\tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 > \underline{Q}_0 | W = w) &= \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1 | W = w) - \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 \leq \underline{Q}_0 | W = w) \\ &= \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) - \min\left\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \underline{F}_{Y_0|W}(\underline{Q}_0|w)\right\} \\ &= \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) \\ &= 0.\end{aligned}$$

Based on (37), we can decompose the left-hand-side term of (36) as below

$$\begin{aligned}\tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)p^{ul}(Y_1, Y_0, w) | W = w) &= \underline{c}\tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq y_0 | W = w) + p_{1|w} \cdot \tilde{\mathbb{P}}(\bar{Q}_1 < Y_1 \leq y_1, Y_0 \leq y_0 | W = w) \\ &= \underline{c}\tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq y_0 | W = w) \\ &= \underline{c} \min\left\{\bar{F}_{Y_1|W}(\bar{Q}_1 | w), \underline{F}_{Y_0|W}(y_0 | w)\right\}.\end{aligned}$$

Note that $\bar{F}_{Y_1|W}(\bar{Q}_1 | w) \geq \underline{F}_{Y_0|W}(\underline{Q}_0 - |w) \geq \underline{F}_{Y_0|W}(y_0 | w)$ by Lemma 7.8 and the condition that $y_0 < \underline{Q}_0$.

We have

$$\underline{c} \min \left\{ \overline{F}_{Y_1|W}(\overline{Q}_1 | w), \underline{F}_{Y_0|W}(y_0 | w) \right\} = \underline{c} \underline{F}_{Y_0|W}(y_0 | w) = p_{1|w} \underline{F}_{Y_0|X,W}(y_0 | 1, w).$$

By Lemma 7.6 (the second set of equivalent results), $\overline{F}_{Y_1|W}(\overline{Q}_1 | w) \geq \underline{F}_{Y_0|W}(y_0 | w)$ also implies

$$\underline{F}_{Y_0|X,W}(y_0 | 1, w) \leq F_{Y|X,W}(\overline{Q}_1 | 1, w) \leq F_{Y|X,W}(y_1 | 1, w),$$

hence we have

$$p_{1|w} \underline{F}_{Y_0|X,W}(y_0 | 1, w) = p_{1|w} \min \left\{ F_{Y|X,W}(y_1 | 1, w), \underline{F}_{Y_0|X,W}(y_0 | 1, w) \right\}.$$

Combining those results then yields (36), as desired.

(Part 1) Case 3: $y_1 < \overline{Q}_1$ and $y_0 \geq \underline{Q}_0$.

Following similar arguments in Case 2, we have the following equality

$$\begin{aligned} & \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{ul}(Y_1, Y_0, w) | W = w) \\ &= \underline{c} \widetilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq \underline{Q}_0 | W = w) + p_{1|w} \cdot \widetilde{\mathbb{P}}(Y_1 \leq y_1, \underline{Q}_0 < Y_0 \leq y_0 | W = w) \\ &= \underline{c} \min \left\{ \overline{F}_{Y_1|W}(y_1 | w), \underline{F}_{Y_0|W}(\underline{Q}_0 | w) \right\} \\ &= \underline{c} \overline{F}_{Y_1|W}(y_1 | w) \\ &= p_{1|w} F_{Y|X,W}(y_1 | 1, w) \\ &= p_{1|w} \min \left\{ F_{Y|X,W}(y_1 | 1, w), \underline{F}_{Y_0|X,W}(y_0 | 1, w) \right\} \end{aligned}$$

where we use (37) in the second equality, the third equality follows by Lemma 7.8, and the condition that $y_1 < \overline{Q}_1$, and the last line holds by Lemma 7.9, where we deduce that

$$F_{Y|X,W}(y_1 | 1, w) \leq F_{Y|X,W}(\overline{Q}_1 | 1, w) \leq \underline{F}_{Y_0|X,W}(\underline{Q}_0 | 1, w) \leq \underline{F}_{Y_0|X,W}(y_0 | 1, w).$$

Therefore, we established (36).

(Part 1) Case 4: $y_1 = \overline{Q}_1$ and $y_0 = \underline{Q}_0$.

We can decompose the LHS probability of (36) as follows:

$$\begin{aligned} & \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{ul}(Y_1, Y_0, w) | W = w) \\ &= \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 = y_1, Y_0 = y_0) p^{ul}(Y_1, Y_0, w) | W = w) + \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 < y_0) p^{ul}(Y_1, Y_0, w) | W = w) \\ &\quad + \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 < y_1, Y_0 \leq y_0) p^{ul}(Y_1, Y_0, w) | W = w) - \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 < y_1, Y_0 < y_0) p^{ul}(Y_1, Y_0, w) | W = w). \\ &= B^{ul} \widetilde{\mathbb{P}}(Y_1 = \overline{Q}_1, Y_0 = \underline{Q}_0 | W = w) \\ &\quad + \lim_{u \nearrow \underline{Q}_0} \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq \overline{Q}_1, Y_0 \leq u | w = w) p^{ul}(Y_1, Y_0, w) | W = w) \\ &\quad + \lim_{v \nearrow \overline{Q}_1} \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq v, Y_0 \leq \underline{Q}_0) p^{ul}(Y_1, Y_0, w) | W = w) \\ &\quad - \lim_{v \nearrow \overline{Q}_1, u \nearrow \underline{Q}_0} \widetilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq v, Y_0 \leq u) p^{ul}(Y_1, Y_0, w) | W = w) \\ &= \widetilde{\mathbb{P}}(Y_1 = \overline{Q}_1, Y_0 = \underline{Q}_0, X = 1 | W = w) + \lim_{u \nearrow \underline{Q}_0} \widetilde{\mathbb{P}}(Y_1 \leq \overline{Q}_1, Y_0 \leq u, X = 1 | W = w) \\ &\quad + \lim_{v \nearrow \overline{Q}_1} \widetilde{\mathbb{P}}(Y_1 \leq v, Y_0 \leq \underline{Q}_0, X = 1 | W = w) - \lim_{v \nearrow \overline{Q}_1, u \nearrow \underline{Q}_0} \widetilde{\mathbb{P}}(Y_1 \leq v, Y_0 \leq u, X = 1 | W = w) \\ &= \widetilde{\mathbb{P}}(Y_1 \leq \overline{Q}_1, Y_0 \leq \underline{Q}_0, X = 1 | W = w). \end{aligned}$$

The second equality holds by the monotone convergence theorem. The third equality holds by the conclusion proved in Case 1–3. The last equality holds by the continuity of probability measure. Thus (36) has been verified.

(Part 1) Case 5: $(y_1, y_0) \geq (\bar{Q}_1, \underline{Q}_0)$.

Given the above results, we have the following derivation:

$$\begin{aligned}
& \tilde{\mathbb{E}} \left(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{ul}(Y_1, Y_0, w) \mid W = w \right) \\
&= \tilde{\mathbb{E}} \left(\mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \underline{Q}_0) p^{ul}(Y_1, Y_0, w) \mid W = w \right) \\
&\quad + \tilde{\mathbb{E}} \left(\bar{c} \left[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) - \mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \underline{Q}_0) \right] \mid W = w \right) \\
&= \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \underline{Q}_0, X = 1 \mid W = w) + \bar{c} \left(\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) - \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \underline{Q}_0) \mid W = w \right) \\
&= \bar{c} \min \left\{ \bar{F}_{Y_1|W}(y_1 \mid w), \underline{F}_{Y_0|W}(y_0 \mid w) \right\} + p_{1|w} \min \left\{ F_{Y|X,W}(\bar{Q}_1 \mid 1, w), \underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) \right\} \\
&\quad - \bar{c} \min \left\{ \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w), \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) \right\} \\
&= \bar{c} \min \left\{ \bar{F}_{Y_1|W}(y_1 \mid w), \underline{F}_{Y_0|W}(y_0 \mid w) \right\} \\
&\quad + [F_{Y|X,W}(\bar{Q}_1 \mid 1, w) p_{1|w} - \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) \bar{c}] \mathbb{1}(\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) \leq \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w)) \\
&\quad + [\underline{F}_{Y_0|X,W}(\underline{Q}_0 \mid 1, w) p_{1|w} - \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w) \bar{c}] \mathbb{1}(\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) > \underline{F}_{Y_0|W}(\underline{Q}_0 \mid w)) \\
&= \bar{c} \min \left\{ \bar{F}_{Y_1|W}(y_1 \mid w), \underline{F}_{Y_0|W}(y_0 \mid w) \right\} - (\bar{c} - p_{1|w}) \\
&= \min \left\{ F_{Y|X,W}(y_1 \mid 1, w) p_{1|w}, \frac{p_{1|w} - \bar{c}}{1 - \bar{c}} + F_{Y|X,W}(y_0 \mid 0, w) \frac{p_{0|w} \bar{c}}{1 - \bar{c}} \right\} \\
&= p_{1|w} \min \left\{ F_{Y|X,W}(y_1 \mid 1, w), \underline{F}_{Y_0|X,W}(y_0 \mid 0, w) \right\},
\end{aligned}$$

where the first equality follows by (37) that (Y_1, Y_0) has no mass on the off-diagonal area, the second equality follows by the result established in Case 4 above, and the fourth equality follows by Lemma 7.6. Thus we have verified (36).

Since \mathbb{R}^2 is partitioned by these 5 cases, we have established that $\tilde{\mathbb{E}}[X \mid Y_1, Y_0, W = w] = p^{ul}(Y_1, Y_0, w)$ almost surely, which concludes the proof of Part 1.

Proof of Part 2: We prove this by constructing a probability distribution $\tilde{\mathbb{P}}$ for (Y_1, Y_0, X) conditional on $W = w$ such that for all $y \in \mathbb{R}$ and $x \in \{0, 1\}$, we have

1. $\tilde{\mathbb{P}}(Y_1 \leq y \mid W = w) = \bar{F}_{Y_1|W}(y \mid w)$ and $\tilde{\mathbb{P}}(Y_0 \leq y \mid W = w) = \bar{F}_{Y_0|W}(y \mid w)$;
2. $\tilde{\mathbb{P}}(X = x \mid W = w) = p_{x|w}$;
3. $\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) = F_{Y|X,W}(Y \mid X, w)$;
4. $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid X = x, W = w) = \max \left\{ \tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = x, W = w) + \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = x, W = w) - 1, 0 \right\}$;
5. $\tilde{\mathbb{P}}(X = 1 \mid Y_1, Y_0, W = w) \in [\underline{c}, \bar{c}]$ for $\tilde{\mathbb{P}}$ -almost surely.

Let

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x \mid W = w) &= x \max \{ F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0 \} p_{1|w} \\
&\quad + (1 - x) \max \{ \bar{F}_{Y_1|X,W}(y_1 \mid 0, w) + F_{Y|X,W}(y_0 \mid 0, w) - 1, 0 \} p_{0|w}.
\end{aligned} \tag{38}$$

By Lemma 4.1, $\bar{F}_{Y_0|X,W}(y_0 \mid 1, w)$ and $\bar{F}_{Y_1|X,W}(y_1 \mid 0, w)$ are cdfs. Also note that $(u, v) \mapsto \max\{u + v - 1, 0\}$ is the counter-monotonic copula. Following Sklar's Theorem, $\tilde{\mathbb{P}}$ is a joint distribution function for (Y_1, Y_0, X) conditional on $W = w$.

Following the same steps as in the proof of Theorem 1, we can show that conditions 1–4 are satisfied because the distribution in (38) is the same as in (21) but for a specific rather than an arbitrary choice of copulas. By Lemma 10, $\tilde{\mathbb{P}}$ leads to the following counter-monotonic joint distribution of (Y_1, Y_0) :

$$\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) = \max\{\bar{F}_{Y_1|W}(y_1 \mid w) + \bar{F}_{Y_0|W}(y_0 \mid w) - 1, 0\}. \quad (39)$$

To show condition 5 holds, and thus complete the proof, we must find a function p^{uu} such that $p^{uu}(Y_1, Y_0, w) = \tilde{\mathbb{E}}[X \mid Y_1, Y_0, W = w]$ and $p^{uu}(Y_1, Y_0, w) \in [\underline{c}, \bar{c}]$ almost surely under $\tilde{\mathbb{P}}$.

First consider $p_{1|w} = \underline{c}$, then we have

$$\bar{F}_{Y_1|X,W}(y_1 \mid 0, w) = \bar{F}_{Y_1|W}(y_1 \mid w) = F_{Y|X,W}(y_1 \mid 1, w)$$

and

$$\bar{F}_{Y_0|X,W}(y_0 \mid 1, w) = \bar{F}_{Y_0|W}(y_0 \mid w) = F_{Y|X,W}(y_0 \mid 0, w).$$

This implies the following derivations:

$$\begin{aligned} \tilde{\mathbb{E}}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) \underline{c} \mid W = w] &= p_{1|w} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) \\ &= p_{1|w} \max\{F_{Y|X,W}(y_1 \mid 1, w) + F_{Y|X,W}(y_0 \mid 0, w) - 1, 0\} \\ &= \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w). \end{aligned}$$

The first line holds by $\underline{c} = p_{1|w}$, the second line holds by (39), and the last line holds by (38). Following Lemma 8, we conclude that $\tilde{\mathbb{P}}(X = 1 \mid Y_1, Y_0, W = w) = \underline{c} \in [\underline{c}, \bar{c}]$ almost surely. The proof of the case where $p_{1|w} = \bar{c}$ is similar and thus omitted.

Next consider $\underline{c} < p_{1|w} < \bar{c}$. Let $p^{uu}(Y_1, Y_0, w) = p^{uu}(Y_1, Y_0, w; B^{uu})$ defined in (24), where

$$B^{uu} = \frac{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0, X = 1 \mid W = w)}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid W = w)}$$

whenever $\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid W = w) > 0$. Set $B^{uu} = p_{1|w}$ otherwise.

We verify that $B^{uu} \in [\underline{c}, \bar{c}]$ if the denominator is nonzero.

First, note that the denominator of B^{uu} can be expanded below

$$\begin{aligned} &\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid W = w) \\ &= \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0 \mid W = w) - \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 < \bar{Q}_0 \mid W = w) \\ &\quad - \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 \leq \bar{Q}_0 \mid W = w) + \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 < \bar{Q}_0 \mid W = w) \\ &= \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1, 0\} - \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) - 1, 0\} \\ &\quad - \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1, 0\} + \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) - 1, 0\}. \end{aligned}$$

where the second equality holds via (39). By Lemma 7.8, we observe that

$$\begin{aligned} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1 &\geq \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} + \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} - 1 = 0 \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) - 1 &\leq \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} + \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} - 1 = 0, \end{aligned}$$

hence this expression simplifies to

$$\begin{aligned}
& \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid W = w) \\
&= \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1 \\
&\quad - \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) - 1, 0\} - \max\{\bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1, 0\} \\
&= \min\{1 - \bar{F}_{Y_0|W}(\bar{Q}_0 - |w), \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w)\} + \min\{1 - \bar{F}_{Y_1|W}(\bar{Q}_1 - |w), \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w)\} - 1.
\end{aligned}$$

Second, we expand the numerator of B^{uu} . We have that

$$\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0, X = 1 \mid W = w) = \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid X = 1, W = w)p_{1|w}$$

and that

$$\begin{aligned}
& \tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid X = 1, W = w) \\
&= \max\{F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1, 0\} \\
&\quad - \max\{F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w) - 1, 0\} \\
&\quad - \max\{F_{Y|X,W}(\bar{Q}_1 - |1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1, 0\} \\
&\quad + \max\{F_{Y|X,W}(\bar{Q}_1 - |1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w) - 1, 0\} \\
&= F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1 - \max\{F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w) - 1, 0\} \\
&\quad - \max\{F_{Y|X,W}(\bar{Q}_1 - |1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1, 0\} + 0 \\
&= \min\{1 - \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w), F_{Y|X,W}(\bar{Q}_1 \mid 1, w)\} + \min\{1 - F_{Y|X,W}(\bar{Q}_1 - |1, w), \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w)\} - 1.
\end{aligned}$$

where the second to last equality follows from Lemma 7.9, where we note that $\bar{\tau}_1 + \underline{\tau}_1 = 1$.

From Lemma 7.11, B^{uu} can take four possible values as follows:

$$\begin{aligned}
B^{uu} &= \frac{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid X = 1, W = w)p_{1|w}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 \mid W = w)} \\
&= \frac{(1 - \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w) - F_{Y|X,W}(\bar{Q}_1 - |1, w)) p_{1|w}}{1 - \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - |w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) \geq 1, \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) \geq 1 \end{array} \right) \quad (40)
\end{aligned}$$

$$+ \frac{(F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 - |1, w)) p_{1|w}}{\bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - |w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) < 1, \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) \geq 1 \end{array} \right) \quad (41)$$

$$+ \frac{(\bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - \bar{F}_{Y_0|X,W}(\bar{Q}_0 - |1, w)) p_{1|w}}{\bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - \bar{F}_{Y_0|W}(\bar{Q}_0 - |w)} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) \geq 1, \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) < 1 \end{array} \right) \quad (42)$$

$$+ \frac{(\bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) + F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - 1) p_{1|w}}{\bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) + \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - 1} \mathbb{1} \left(\begin{array}{l} \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - |w) < 1, \\ \bar{F}_{Y_1|W}(\bar{Q}_1 - |w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) < 1 \end{array} \right) \quad (43)$$

As shown in Lemma 7.3, terms (41) and (42) lie in $[\underline{c}, \bar{c}]$.

Next we examine the term (40), which can be written as below

$$\begin{aligned}
& \frac{(1 - \bar{F}_{Y_0|X,W}(\bar{Q}_0 - | 1, w) - F_{Y|X,W}(\bar{Q}_1 - | 1, w)) p_{1|w}}{1 - \bar{F}_{Y_0|W}(\bar{Q}_0 - | w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - | w)} \\
&= \frac{p_{1|w} \left(1 - F_{Y|X,W}(\bar{Q}_0 - | 0, w) \frac{p_{0|w} \bar{c}}{p_{1|w}(1-\bar{c})} - F_{Y|X,W}(\bar{Q}_1 - | 1, w) \right)}{1 - F_{Y|X,W}(\bar{Q}_0 - | 0, w) \frac{p_{0|w}}{1-\bar{c}} - F_{Y|X,W}(\bar{Q}_1 - | 1, w) \frac{p_{1|w}}{\underline{c}}} \\
&= \underline{c} + \frac{p_{1|w} - \underline{c} - F_{Y|X,W}(\bar{Q}_0 - | 0, w) \frac{p_{0|w}(\bar{c}-\underline{c})}{1-\bar{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \\
&\geq \underline{c} + \frac{p_{1|w} - \underline{c} - \bar{\tau}_0 \frac{p_{0|w}(\bar{c}-\underline{c})}{1-\bar{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \\
&= \underline{c},
\end{aligned}$$

where the inequality follows by Lemma 6.3 that $F_{Y|X,W}(\bar{Q}_0 - | 0, w) = F_{Y|X,W}(Q_{Y|X,W}(\bar{\tau}_0 | 0, w) - | 0, w) \leq \bar{\tau}_0$. Also note that

$$\begin{aligned}
\frac{(1 - \bar{F}_{Y_0|X,W}(\bar{Q}_0 - | 1, w) - F_{Y|X,W}(\bar{Q}_1 - | 1, w)) p_{1|w}}{1 - \bar{F}_{Y_0|W}(\bar{Q}_0 - | w) - \bar{F}_{Y_1|W}(\bar{Q}_1 - | w)} &= \bar{c} + \frac{(p_{1|w} - \bar{c}) + F_{Y|X,W}(\bar{Q}_1 - | 1, w) \frac{p_{1|w}(\bar{c}-\underline{c})}{\underline{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \\
&\leq \bar{c} + \frac{(p_{1|w} - \bar{c}) + \bar{\tau}_1 \frac{p_{1|w}(\bar{c}-\underline{c})}{\underline{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \\
&= \bar{c},
\end{aligned}$$

where the inequality follows by Lemma 6.3 that $F_{Y|X,W}(\bar{Q}_1 - | 1, w) = F_{Y|X,W}(Q_{Y|X,W}(\bar{\tau}_1 | 1, w) - | 1, w) \leq \bar{\tau}_1$. Then we have shown that the term (40) is bounded between \underline{c} and \bar{c} .

Then consider the term (43). Following the same arguments, we have

$$\begin{aligned}
\frac{(\bar{F}_{Y_0|X,W}(\bar{Q}_0 | 1, w) + F_{Y|X,W}(\bar{Q}_1 | 1, w) - 1) p_{1|w}}{\bar{F}_{Y_0|W}(\bar{Q}_0 | w) + \bar{F}_{Y_1|W}(\bar{Q}_1 | w) - 1} &= \underline{c} + \frac{\frac{p_{1|w}(\bar{c}-\underline{c})}{\bar{c}} F_{Y|X,W}(\bar{Q}_1 | 1, w) - \frac{(\bar{c}-p_{1|w})\underline{c}}{\bar{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \geq \underline{c} \\
\frac{(\bar{F}_{Y_0|X,W}(\bar{Q}_0 | 1, w) + F_{Y|X,W}(\bar{Q}_1 | 1, w) - 1) p_{1|w}}{\bar{F}_{Y_0|W}(\bar{Q}_0 | w) + \bar{F}_{Y_1|W}(\bar{Q}_1 | w) - 1} &= \bar{c} + \frac{\frac{(p_{1|w}-\underline{c})(1-\bar{c})}{1-\underline{c}} - F_{Y|X,W}(\bar{Q}_0 | 0, w) \frac{p_{0|w}(\bar{c}-\underline{c})}{1-\underline{c}}}{\tilde{\mathbb{P}}(Y_1 = \bar{Q}_1, Y_0 = \bar{Q}_0 | W = w)} \leq \bar{c},
\end{aligned}$$

where inequalities follow by Lemma 6.2 that $F_{Y|X,W}(\bar{Q}_x | x, w) \geq \bar{\tau}_x$ for $x = 0, 1$. So we have shown that all four terms (40)–(43) are bounded within $[\underline{c}, \bar{c}]$, thus concluding $B^{uu} \in [\underline{c}, \bar{c}]$, which then establishes $p^{uu}(Y_1, Y_0, w) \in [\underline{c}, \bar{c}]$ almost surely.

To finish this proof, we demonstrate that $\tilde{\mathbb{E}}[X | Y_1, Y_0, W = w] = p^{uu}(Y_1, Y_0, w)$ almost surely. To do so, we use Lemma 8 and show that

$$\begin{aligned}
\tilde{\mathbb{E}}[\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{uu}(Y_1, Y_0, w) | W = w] &= \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 | W = w) \\
&= p_{1|w} \max\{F_{Y|X,W}(y_1 | 1, w) + \bar{F}_{Y_0|X,W}(y_0 | 1, w) - 1, 0\}
\end{aligned} \tag{44}$$

for all $(y_1, y_0) \in \mathbb{R}^2$. To complete the proof, we break this up into different cases.

(Part 2) Case 1: $y_1 < \bar{Q}_1$ and $y_0 < \bar{Q}_0$.

First, note that the joint cdf from (39) implies

$$\tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 < \bar{Q}_0 | W = w) = \tilde{\mathbb{P}}(\bar{Q}_1 < Y_1, \bar{Q}_0 < Y_0 | W = w) = 0. \tag{45}$$

These equalities can be verified by the arguments below:

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 < \bar{Q}_0 \mid W = w) &= \max \{ \bar{F}_{Y_1|W}(\bar{Q}_1 - | w) + \bar{F}_{Y_0|W}(\bar{Q}_0 - | w) - 1, 0 \} \\
&\leq \max \left\{ \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} + \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} - 1, 0 \right\} \\
&= 0,
\end{aligned}$$

where the inequality follows by Lemma 7.8, and similarly,

$$\begin{aligned}
\tilde{\mathbb{P}}(\bar{Q}_1 < Y_1, \bar{Q}_0 < Y_0 \mid W = w) &= 1 - \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1 \mid W = w) - \tilde{\mathbb{P}}(Y_0 \leq \bar{Q}_0 \mid W = w) \\
&\quad + \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0 \mid W = w) \\
&= 1 - \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) - \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) \\
&\quad + \max \{ \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) - 1, 0 \} \\
&= 1 - \min \{ 1, \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w) + \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w) \} \\
&\leq 1 - \min \left\{ 1, \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} + \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} \right\} \\
&= 0,
\end{aligned}$$

where the inequality follows by Lemma 7.8.

On the one hand, (45) implies

$$\begin{aligned}
\tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{uu}(Y_1, Y_0, w) \mid W = w) &= p_{1|w} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid W = w) \\
&\leq p_{1|w} \tilde{\mathbb{P}}(Y_1 < \bar{Q}_1, Y_0 < \bar{Q}_0 \mid W = w) \\
&= 0.
\end{aligned}$$

This shows $\tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{uu}(Y_1, Y_0, w) \mid W = w) = 0$ due to the construction that $p^{uu}(Y_1, Y_0, w)$ is non-negative. On the other hand,

$$\begin{aligned}
\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w) &= p_{1|w} \max \{ F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0 \} \\
&\leq p_{1|w} \max \{ F_{Y|X,W}(\bar{Q}_1 - \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 - \mid 1, w) - 1, 0 \} \\
&\leq p_{1|w} \max \{ \bar{\tau}_1 + \underline{\tau}_1 - 1, 0 \} \\
&= 0,
\end{aligned}$$

where the second inequality follows by Lemma 7.9. This implies $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w) = 0$, thus establishing (44), as desired.

(Part 2) Case 2: $y_1 \geq \bar{Q}_1, y_0 < \bar{Q}_0$.

First, note that (45) implies that

$$\begin{aligned}
&\tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{uu}(Y_1, Y_0, w) \mid W = w) \\
&= \tilde{\mathbb{E}}(\mathbb{1}(\bar{Q}_1 \leq Y_1 \leq y_1, Y_0 \leq y_0) \bar{c} \mid W = w) + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 < \bar{Q}_1, Y_0 \leq y_0) p_{1|w} \mid W = w) \\
&= \tilde{\mathbb{E}}(\mathbb{1}(\bar{Q}_1 \leq Y_1 \leq y_1, Y_0 \leq y_0) \bar{c} \mid W = w) + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 < \bar{Q}_1, Y_0 \leq y_0) \bar{c} \mid W = w) \\
&= \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) \bar{c} \mid W = w),
\end{aligned}$$

where the second equality follows by the fact that $\tilde{\mathbb{P}}$ takes no mass on $\{Y_1 < \bar{Q}_1, Y_0 \leq y_0\} \subseteq \{Y_1 < \bar{Q}_1, Y_0 <$

$\bar{Q}_0\}$ by (45). Next we expand the last expression

$$\begin{aligned}
& \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)\bar{c}|W = w) \\
&= \bar{c}\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 | W = w) \\
&= \bar{c} \max \left\{ \bar{F}_{Y_1|W}(y_1 | w) + \bar{F}_{Y_0|W}(y_0 | w) - 1, 0 \right\} \\
&= \bar{c} \max \left\{ \frac{\bar{c} - p_{1|w}}{\bar{c}} + F_{Y|X,W}(y_1 | 1, w) \frac{p_{1|w}}{\bar{c}} + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}}{1 - \bar{c}} - 1, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + \frac{\bar{c} - p_{1|w}}{p_{1|w}} + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1 - \bar{c})} - \frac{\bar{c}}{p_{1|w}}, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}\bar{c}}{p_{1|w}(1 - \bar{c})} - 1, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + \bar{F}_{Y_0|X,W}(y_0 | 1, w) - 1, 0 \right\},
\end{aligned}$$

thus establishing (44), as desired.

(Part 2) Case 3: $y_1 < \bar{Q}_1, y_0 \geq \bar{Q}_0$.

Similar to the proof of case 2 above, we have

$$\tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)p^{uu}(Y_1, Y_0, w) | W = w) = \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)\underline{c}|W = w)$$

due to (45). Next we expand the expression on the right hand side.

$$\begin{aligned}
& \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0)\underline{c}|W = w) \\
&= \underline{c} \max \left\{ \bar{F}_{Y_1|W}(y_1 | w) + \bar{F}_{Y_0|W}(y_0 | w) - 1, 0 \right\} \\
&= \underline{c} \max \left\{ F_{Y|X,W}(y_1 | 1, w) \frac{p_{1|w}}{\underline{c}} + \frac{p_{1|w} - \underline{c}}{1 - \underline{c}} + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}}{1 - \underline{c}} - 1, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + \frac{(p_{1|w} - \underline{c})\underline{c}}{p_{1|w}(1 - \underline{c})} + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1 - \underline{c})} - \frac{\underline{c}}{p_{1|w}}, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + \frac{p_{1|w} - \underline{c}}{p_{1|w}(1 - \underline{c})} + F_{Y|X,W}(y_0|0, w) \frac{p_{0|w}\underline{c}}{p_{1|w}(1 - \underline{c})} - 1, 0 \right\} \\
&= p_{1|w} \max \left\{ F_{Y|X,W}(y_1 | 1, w) + \bar{F}_{Y_0|X,W}(y_0 | 1, w) - 1, 0 \right\},
\end{aligned}$$

thus establishing (44), as desired.

(Part 2) Case 4: $y_1 = \bar{Q}_1, y_0 = \bar{Q}_0$.

Note that the equality (44) can be established following the same arguments from the proof of Part 1, case 4. Once the results are established for cases 1–3, the equality (44) holds for $y_1 = \bar{Q}_1, y_0 = \bar{Q}_0$ by applying monotone convergence theorem and continuity of measure. To this end, the proof is omitted.

(Part 2) Case 5: $(y_1, y_0) \geq (\bar{Q}_1, \bar{Q}_0)$.

We start by noting that

$$\begin{aligned}
& \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq y_1, Y_0 \leq y_0) p^{uu}(Y_1, Y_0, w) \mid W = w) \\
&= p_{1|w} \tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1], Y_0 \in (\bar{Q}_0, y_0] \mid W = w) + \underline{c} \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \in (\bar{Q}_0, y_0] \mid W = w) \\
&\quad + \bar{c} \tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1], Y_0 \leq \bar{Q}_0 \mid W = w) + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0) p^{uu}(Y_1, Y_0, w) \mid W = w) \\
&= \underline{c} \left(\tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \in (\bar{Q}_0, y_0] \mid W = w) + \tilde{\mathbb{P}}(Y_1 > \bar{Q}_1, Y_0 \in (\bar{Q}_0, y_0] \mid W = w) \right) \\
&\quad + \bar{c} \left(\tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1], Y_0 \leq \bar{Q}_0 \mid W = w) + \tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1], Y_0 > \bar{Q}_0 \mid W = w) \right) \\
&\quad + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0) p^{uu}(Y_1, Y_0, w) \mid W = w) \\
&= \underline{c} \tilde{\mathbb{P}}(Y_0 \leq (\bar{Q}_0, y_0] \mid W = w) + \bar{c} \tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1] \mid W = w) + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0) p^{uu}(Y_1, Y_0, w) \mid W = w),
\end{aligned}$$

where the second equality follows by (45) that $\tilde{\mathbb{P}}$ takes no mass on diagonal area. Next we expand the last line.

$$\begin{aligned}
& \underline{c} \tilde{\mathbb{P}}(Y_0 \leq (\bar{Q}_0, y_0] \mid W = w) + \bar{c} \tilde{\mathbb{P}}(Y_1 \in (\bar{Q}_1, y_1] \mid W = w) + \tilde{\mathbb{E}}(\mathbb{1}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0) p^{uu}(Y_1, Y_0) \mid W = w) \\
&= \underline{c} [\bar{F}_{Y_0|W}(y_0 \mid w) - \bar{F}_{Y_0|W}(\bar{Q}_0 \mid w)] + \bar{c} [\bar{F}_{Y_1|W}(y_1 \mid w) - \bar{F}_{Y_1|W}(\bar{Q}_1 \mid w)] \\
&\quad + \tilde{\mathbb{P}}(Y_1 \leq \bar{Q}_1, Y_0 \leq \bar{Q}_0, X = 1 \mid W = w) \\
&= \frac{p_{0|w} \underline{c}}{1 - \underline{c}} (F_{Y|X,W}(y_0 \mid 0, w) - F_{Y|X,W}(\bar{Q}_0 \mid 0, w)) + p_{1|w} (F_{Y|X,W}(y_1 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 \mid 1, w)) \\
&\quad + p_{1|w} \max \{ F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1, 0 \} \\
&= \frac{p_{0|w} \underline{c}}{1 - \underline{c}} (F_{Y|X,W}(y_0 \mid 0, w) - F_{Y|X,W}(\bar{Q}_0 \mid 0, w)) + p_{1|w} (F_{Y|X,W}(y_1 \mid 1, w) - F_{Y|X,W}(\bar{Q}_1 \mid 1, w)) \\
&\quad + p_{1|w} (F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1) \\
&= p_{1|w} \left[F_{Y|X,W}(y_1 \mid 1, w) + \frac{p_{1|w} - \underline{c}}{p_{1|w}(1 - \underline{c})} + \frac{p_{0|w} \underline{c}}{p_{1|w}(1 - \underline{c})} F_{Y|X,W}(y_0 \mid 0, w) - 1 \right] \\
&= p_{1|w} [F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1] \\
&= p_{1|w} \max \{ F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1, 0 \}.
\end{aligned}$$

The third and the last equality hold by the following derivation

$$\begin{aligned}
F_{Y|X,W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) - 1 &\geq F_{Y|X,W}(\bar{Q}_1 \mid 1, w) + \bar{F}_{Y_0|X,W}(\bar{Q}_0 \mid 1, w) - 1 \\
&\geq \bar{\tau}_1 + \underline{\tau}_1 - 1 \\
&= 0,
\end{aligned}$$

where the second inequality follows by Lemma 7.9. Hence we have established (44), as desired.

Since \mathbb{R}^2 is partitioned by these 5 cases, we haven shown that $\tilde{\mathbb{E}}[X \mid Y_1, Y_0, W = w] = p^{uu}(Y_1, Y_0, w)$ almost surely, which concludes the proof of Part 2.

Proof of Part 3: One can show that $(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, \bar{C}_{1,0|X,W})$ can be achieved by the joint distribution of (Y_1, Y_0, X) conditional on $W = w$ constructed as below:

$$\begin{aligned}
& \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w) \\
&= x \min \{ F_{Y|X,W}(y_1 \mid 1, w), \bar{F}_{Y_0|X,W}(y_0 \mid 1, w) \} p_{1|w} + (1 - x) \min \{ \underline{F}_{Y_1|X,W}(y_1 \mid 0, w), F_{Y|X,W}(y_0 \mid 0, w) \} p_{0|w}.
\end{aligned}$$

It can be verified that this joint distribution satisfies the following 5 conditions: for all $y \in \mathbb{R}$ and $x \in \{0, 1\}$,

1. $\tilde{\mathbb{P}}(Y_1 \leq y \mid W = w) = \underline{F}_{Y_1|W}(y \mid w)$ and $\tilde{\mathbb{P}}(Y_0 \leq y \mid W = w) = \bar{F}_{Y_0}(y \mid w)$;
2. $\tilde{\mathbb{P}}(X = x \mid W = w) = p_{x|w}$;

3. $\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) = F_{Y \mid X, W}(Y \mid X, w)$;
4. $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid X = x, W = w) = \min \left\{ \tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = x, W = w), \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = x, W = w) \right\}$;
5. $\tilde{\mathbb{E}}(X \mid Y_1, Y_0, W = w) = p^{lu}(Y_1, Y_0, w; B^{lu}) \in [\underline{c}, \bar{c}]$, for $\tilde{\mathbb{P}}$ -almost surely with

$$B^{lu} = \frac{\tilde{\mathbb{P}}(Y_1 = y_1, Y_0 = y_0, X = 1 \mid W = w)}{\tilde{\mathbb{P}}(Y_1 = y_1, Y_0 = y_0 \mid W = w)}.$$

The arguments are similar to the proof of Part 1 and thus omitted.

Proof of Part 4: One can show that $(\underline{F}_{Y_1 \mid W}, \underline{F}_{Y_0 \mid W}, \underline{C}_{1,0 \mid X, W})$ can be achieved by the joint distribution of (Y_1, Y_0, X) conditional on $W = w$ constructed as below:

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1 \mid W = w) &= x \max \left\{ F_{Y \mid X, W}(y_1 \mid 1, w) + \underline{F}_{Y_0 \mid X, W}(y_0 \mid 1, w) - 1, 0 \right\} p_{1 \mid w} \\ &\quad + (1 - x) \max \left\{ \underline{F}_{Y_1 \mid X, W}(y_1 \mid 0, w) + F_{Y \mid X, W}(y_0 \mid 0, w) - 1, 0 \right\} p_{0 \mid w}. \end{aligned}$$

It can be verified that this joint distribution satisfies the following four conditions: for all $y \in \mathbb{R}$ and $x \in \{0, 1\}$,

1. $\tilde{\mathbb{P}}(Y_1 \leq y \mid W = w) = \underline{F}_{Y_1 \mid W}(y \mid w)$ and $\tilde{\mathbb{P}}(Y_0 \leq y \mid W = w) = \underline{F}_{Y_0 \mid W}(y \mid w)$;
2. $\tilde{\mathbb{P}}(X = x \mid W = w) = p_{x \mid w}$;
3. $\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) = F_{Y \mid X, W}(Y \mid X, w)$;
4. $\tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0 \mid X = x, W = w) = \max \left\{ \tilde{\mathbb{P}}(Y_1 \leq y_1 \mid X = x, W = w) + \tilde{\mathbb{P}}(Y_0 \leq y_0 \mid X = x, W = w) - 1, 0 \right\}$;
5. $\tilde{\mathbb{E}}(X \mid Y_1, Y_0, W = w) = p^{lu}(Y_1, Y_0, w; B^{lu}) \in [\underline{c}, \bar{c}]$, for $\tilde{\mathbb{P}}$ -almost surely with

$$B^{lu} = \frac{\tilde{\mathbb{P}}(Y_1 = y_1, Y_0 = y_0, X = 1 \mid W = w)}{\tilde{\mathbb{P}}(Y_1 = y_1, Y_0 = y_0 \mid W = w)}.$$

The arguments are similar to the proof of Part 2 and thus omitted. \square

Suppose $\{F_{Y_1, Y_0, X}^k\}_{k=1}^K$ is a set of valid cdfs of (Y_1, Y_0, X) whose support of X is $\{0, 1\}$. Consider a mixture of cdfs defined as

$$F_{Y_1, Y_0, X}^{\text{mix}}(y_1, y_0, x) = \sum_{k=1}^K a_k F_{Y_1, Y_0, X}^k(y_1, y_0, x) \quad (46)$$

where $a_k \in [0, 1]$ for all $k \in \{1, \dots, K\}$, $\sum_{k=1}^K a_k = 1$, and $(y_1, y_0, x) \in \mathbb{R}^2 \times \{0, 1\}$.

Let $p^k(Y_1, Y_0) := \mathbb{E}^k[X \mid Y_1, Y_0]$, where \mathbb{E}^k denotes the expectation under cdf $F_{Y_1, Y_0, X}^k$ for $k \in \{1, \dots, K, \text{mix}\}$. The next lemma will be used to show that mixtures of distributions satisfying joint c -dependence also satisfy joint c -dependence.

Lemma 12. There exists a sequence of Borel measurable function $\{f_k(\cdot, \cdot)\}_{k=1}^K : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that

1. $f_k(Y_1, Y_0) \in [0, 1]$ for each $k \in \{1, \dots, K\}$, and $\sum_{k=1}^K f_k(Y_1, Y_0) = 1$,
2. $p^{\text{mix}}(Y_1, Y_0) = \sum_{k=1}^K f_k(Y_1, Y_0) p^k(Y_1, Y_0)$,

almost surely under the distribution $F_{Y_1, Y_0, X}^{\text{mix}}$.

Proof of Lemma 12. Let \mathbb{P}^k denote the probability taken under cdf $F_{Y_1, Y_0, X}^k$ for $k \in \{1, \dots, K, \text{mix}\}$. Then it follows by the definition of conditional probability that

$$\begin{aligned} \mathbb{P}^{\text{mix}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1) &= \mathbb{E}^{\text{mix}} [\mathbb{1}[Y_1 \leq y_1, Y_0 \leq y_0] p^{\text{mix}}(Y_1, Y_0)] \\ &= \int_{u \leq y_1, v \leq y_0} p^{\text{mix}}(u, v) dF_{Y_1, Y_0}^{\text{mix}}, \end{aligned}$$

where the last line denotes the Lebesgue-Stieltjes integral with respect to the cdf $F_{Y_1, Y_0}^{\text{mix}}$.

Likewise, for each $k \in \{1, \dots, K\}$, we have

$$\mathbb{P}^k(Y_1 \leq y_1, Y_0 \leq y_0, X = 1) = \int_{u \leq y_1, v \leq y_0} p^k(u, v) dF_{Y_1, Y_0}^k.$$

Since (46) implies that

$$\mathbb{P}^{\text{mix}}(Y_1 \leq y_1, Y_0 \leq y_0, X = 1) = \sum_{k=1}^K a_k \mathbb{P}^k(Y_1 \leq y_1, Y_0 \leq y_0, X = 1),$$

we have

$$\int_{w \leq y_1, v \leq y_0} p^{\text{mix}}(w, v) dF_{Y_1, Y_0}^{\text{mix}} = \sum_{k=1}^K a_k \int_{u \leq y_1, v \leq y_0} p^k(u, v) dF_{Y_1, Y_0}^k. \quad (47)$$

Following the Carathéodory extension theorem (e.g., Ash and Doléans-Dade (2000, Theorem 1.4.9)), there exists unique Lebesgue-Stieltjes measures ν^{mix} and $\{\nu^k\}_{k=1}^K$ defined on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ that are consistent with cdfs $F_{Y_1, Y_0}^{\text{mix}}$ and $\{F_{Y_1, Y_0}^k\}_{k=1}^K$, respectively. Combined with (46), this implies

$$\nu^{\text{mix}}(A) = \sum_{k=1}^K a_k \nu^k(A), \quad \text{for all } A \in \mathcal{A} := \{(-\infty, y_0] \times (-\infty, y_1] : (y_0, y_1) \in \mathbb{R}^2\}. \quad (48)$$

It can be seen that \mathcal{A} is a π -system, $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^2)$, and the class of sets satisfying (48) constitutes a λ -system. Following from $\pi - \lambda$ Theorem,

$$\nu^{\text{mix}}(A) = \sum_{k=1}^K a_k \nu^k(A), \quad \text{for all } A \in \mathcal{B}(\mathbb{R}^2). \quad (49)$$

From the above identity (49), we note that $a_k \nu^k \ll \nu^{\text{mix}}$ for all $k \in \{1, \dots, K\}$. By Radon-Nikodym Theorem (e.g., Royden and Fitzpatrick (2010, p.386, Problem 54.1)), there exist nonnegative Borel measurable functions $d(a_k \nu^k)/d\nu^{\text{mix}}$ such that the following equalities hold for all Borel sets $A \in \mathcal{B}(\mathbb{R}^2)$ and for each $k \in \{1, \dots, K\}$:

$$\int_A p^k(u, v) \frac{d(a_k \nu^k)}{d\nu^{\text{mix}}} d\nu^{\text{mix}} = \int_A p^k(u, v) d(a_k \nu^k) = \int_A a_k p^k(u, v) d\nu^k.$$

Taking $A = (-\infty, y_1] \times (-\infty, y_0]$ and combining these equalities across $k \in \{1, \dots, K\}$ then gives

$$\begin{aligned} \int_{u \leq y_1, v \leq y_0} \sum_{k=1}^K p^k(u, v) \frac{d(a_k \nu^k)}{d\nu^{\text{mix}}} d\nu^{\text{mix}} &= \sum_{k=1}^K a_k \int_{u \leq y_1, v \leq y_0} p^k(u, v) d\nu^k \\ &= \int_{u \leq y_1, v \leq y_0} p^{\text{mix}}(u, v) d\nu^{\text{mix}} \end{aligned}$$

The second equality holds by (47). Since \mathcal{A} constitutes a π -system, and \mathbb{R}^2 can be written as a countable union of elements in the class \mathcal{A} , applying Billingsley (1995, Theorem 16.10.(iii)) then leads to the following

equality:

$$p^{\text{mix}}(u, v) = \sum_{k=1}^K p^k(u, v) \frac{d(a_k \nu^k)}{d\nu^{\text{mix}}}(u, v)$$

almost surely with respect to the measure ν^{mix} on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. Replacing (u, v) with random variables (Y_1, Y_0) , the above equality is equivalent to

$$p^{\text{mix}}(Y_1, Y_0) = \sum_{k=1}^K p^k(Y_1, Y_0) \frac{d(a_k \nu^k)}{d\nu^{\text{mix}}}(Y_1, Y_0)$$

almost surely under the distribution $F_{Y_1, Y_0}^{\text{mix}}$. Next we show that the weights add up to one and they are non-negative.

Following from (49), the conclusion in Royden and Fitzpatrick (2010, p.386, Problem 54.2) implies

$$\sum_{k=1}^K \frac{d(a_k \nu^k)}{d\nu^{\text{mix}}} = \frac{d\nu^{\text{mix}}}{d\nu^{\text{mix}}} = 1, \quad \text{almost surely-}\nu^{\text{mix}}.$$

By definition, Radon-Nikodym derivatives are nonnegative. So we have

$$\frac{d(a_k \nu^k)}{d\nu^{\text{mix}}} = 1 - \sum_{j \neq k} \frac{d(a_j \nu^j)}{d\nu^{\text{mix}}} \in [0, 1], \quad \text{almost surely-}\nu^{\text{mix}}.$$

Therefore, we have shown that $p^{\text{mix}}(Y_1, Y_0)$ is a convex combination of $\{p^k(Y_1, Y_0)\}_{k=1}^K$ almost surely, with weights $d(a_k \nu^k)/d\nu^{\text{mix}}$ being a non-negative measurable function of (Y_1, Y_0) , as desired. \square

Proof of Theorem 2. Fix a $w \in \text{supp}(W)$ and $(\varepsilon, \gamma) \in [0, 1]^2$, we prove this by constructing a probability distribution $\tilde{\mathbb{P}}$ for (Y_1, Y_0, X) conditional on $W = w$ such that for all $y \in \mathbb{R}$ and $x \in \{0, 1\}$, the following conditions hold

1. $\tilde{\mathbb{P}}(Y_1 \leq y \mid W = w) = \varepsilon F_{Y_1|W}(y \mid w) + (1 - \varepsilon) \bar{F}_{Y_1|W}(y \mid w)$ and
 $\tilde{\mathbb{P}}(Y_0 \leq y \mid W = w) = \gamma F_{Y_0|W}(y \mid w) + (1 - \gamma) \bar{F}_{Y_0|W}(y \mid w);$
2. $\tilde{\mathbb{P}}(X = x \mid W = w) = p_{x|w};$
3. $\tilde{\mathbb{P}}(Y_x \leq y \mid X = x, W = w) = F_{Y|X, W}(Y \mid X, w);$
4. $\tilde{\mathbb{P}}(X = 1 \mid Y_1, Y_0, W = w) \in [\underline{c}, \bar{c}]$ for $\tilde{\mathbb{P}}$ -almost surely.

Compared to the proof of Theorem 1, we remove the requirement that the copulas between Y_1 and Y_0 conditional on (X, W) can be arbitrary.

As in Lemma 11, we have constructed the following four joint (conditional) cdfs of $(Y_1, Y_0, X) \mid W = w$:

$$\begin{aligned} F^{ul}(y_1, y_0, x|w) &= x \min \left\{ F_{Y|X, W}(y_1 \mid 1, w), \underline{F}_{Y_0|X}(y_0 \mid 1) \right\} p_{1|w} \\ &\quad + (1 - x) \min \left\{ \bar{F}_{Y_1|X, W}(y_1 \mid 0, w), F_{Y|X, W}(y_0 \mid 0, w) \right\} p_{0|w} \\ F^{lu}(y_1, y_0, x|w) &= x \min \left\{ F_{Y|X, W}(y_1 \mid 1), \bar{F}_{Y_0|X, W}(y_0 \mid 1, w) \right\} p_{1|w} \\ &\quad + (1 - x) \min \left\{ \underline{F}_{Y_1|X, W}(y_1 \mid 0, w), F_{Y|X, W}(y_0 \mid 0, w) \right\} p_{0|w} \\ F^{uu}(y_1, y_0, x|w) &= x \max \left\{ F_{Y|X, W}(y_1 \mid 1, w) + \bar{F}_{Y_0|X, W}(y_0 \mid 1, w) - 1, 0 \right\} p_{1|w} \\ &\quad + (1 - x) \min \left\{ \bar{F}_{Y_1|X, W}(y_1 \mid 0, w) + F_{Y|X, W}(y_0 \mid 0, w) - 1, 0 \right\} p_{0|w} \\ F^{ll}(y_1, y_0, x|w) &= x \max \left\{ F_{Y|X, W}(y_1 \mid 1, w) + \underline{F}_{Y_0|X, W}(y_0 \mid 1, w) - 1, 0 \right\} p_{1|w} \\ &\quad + (1 - x) \min \left\{ \underline{F}_{Y_1|X, W}(y_1 \mid 0, w) + F_{Y|X, W}(y_0 \mid 0, w) - 1, 0 \right\} p_{0|w}. \end{aligned}$$

Let the joint distribution of $(Y_1, Y_0, X) | W = w$ be defined as below:

$$\begin{aligned} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, X = x | W = w) &= \gamma [\varepsilon F^{ll}(y_1, y_0, x|w) + (1 - \varepsilon) F^{ul}(y_1, y_0, x|w)] \\ &\quad + (1 - \gamma) [\varepsilon F^{lu}(y_1, y_0, x|w) + (1 - \varepsilon) F^{uu}(y_1, y_0, x|w)]. \end{aligned} \quad (50)$$

As shown in Lemma 11, the functions F^{ll} , F^{ul} , F^{lu} , and F^{uu} are valid cdfs, hence their convex combination (50) also yields a valid cdf. Next we verify that this cdf satisfies conditions 1-4 listed above, thus concluding that $(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \overline{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \overline{F}_{Y_0|W})$ is in the identified set.

Verifying Condition 1: For $y \in \mathbb{R}$, we have

$$\begin{aligned} &\tilde{\mathbb{P}}(Y_1 \leq y | W = w) \\ &= \lim_{y_0 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y, Y_0 \leq y_0, X = 1 | W = w) + \lim_{y_0 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y, Y_0 \leq y_0, X = 0 | W = w) \\ &= \sum_{x=0,1} \gamma \left[\varepsilon \lim_{y_0 \rightarrow +\infty} F^{ll}(y, y_0, x|w) + (1 - \varepsilon) \lim_{y_0 \rightarrow +\infty} F^{ul}(y, y_0, x|w) \right] \\ &\quad + \sum_{x=0,1} (1 - \gamma) \left[\varepsilon \lim_{y_0 \rightarrow +\infty} F^{lu}(y, y_0, x|w) + (1 - \varepsilon) \lim_{y_0 \rightarrow +\infty} F^{uu}(y, y_0, x|w) \right] \\ &= \gamma \left[\varepsilon \underline{F}_{Y_1|W}(y | w) + (1 - \varepsilon) \overline{F}_{Y_1|W}(y | w) \right] + (1 - \gamma) \left[\varepsilon \underline{F}_{Y_1|W}(y | w) + (1 - \varepsilon) \overline{F}_{Y_1|W}(y | w) \right] \\ &= \varepsilon \underline{F}_{Y_1|W}(y | w) + (1 - \varepsilon) \overline{F}_{Y_1|W}(y | w), \end{aligned}$$

where the third equality uses the conclusion from condition 1 in the proof of Lemma 11.

Likewise,

$$\begin{aligned} &\tilde{\mathbb{P}}(Y_0 \leq y | W = w) \\ &= \lim_{y_1 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y, X = 1 | W = w) + \lim_{y_1 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y, X = 0 | W = w) \\ &= \sum_{x=0,1} \gamma \left[\varepsilon \lim_{y_1 \rightarrow +\infty} F^{ll}(y_1, y, x|w) + (1 - \varepsilon) \lim_{y_1 \rightarrow +\infty} F^{ul}(y_1, y, x|w) \right] \\ &\quad + \sum_{x=0,1} (1 - \gamma) \left[\varepsilon \lim_{y_1 \rightarrow +\infty} F^{lu}(y_1, y, x|w) + (1 - \varepsilon) \lim_{y_1 \rightarrow +\infty} F^{uu}(y_1, y, x|w) \right] \\ &= \gamma \left[\varepsilon \underline{F}_{Y_0|W}(y | w) + (1 - \varepsilon) \overline{F}_{Y_0|W}(y | w) \right] + (1 - \gamma) \left[\varepsilon \overline{F}_{Y_0|W}(y | w) + (1 - \varepsilon) \underline{F}_{Y_0|W}(y | w) \right] \\ &= \gamma \underline{F}_{Y_0|W}(y | w) + (1 - \gamma) \overline{F}_{Y_0|W}(y | w). \end{aligned}$$

Verifying Condition 2: For $x \in \{0, 1\}$, we have that

$$\begin{aligned} \tilde{\mathbb{P}}(X = x | W = w) &= \lim_{y_0, y_1 \rightarrow +\infty} \tilde{\mathbb{P}}(Y_1 \leq y_1, Y_0 \leq y_0, x | W = w) \\ &= \gamma \left[\varepsilon \lim_{y_0, y_1 \rightarrow +\infty} F^{ll}(y_1, y_0, x|w) + (1 - \varepsilon) \lim_{y_0, y_1 \rightarrow +\infty} F^{ul}(y_1, y_0, x|w) \right] \\ &\quad + (1 - \gamma) \left[\varepsilon \lim_{y_0, y_1 \rightarrow +\infty} F^{lu}(y_1, y_0, x|w) + (1 - \varepsilon) \lim_{y_0, y_1 \rightarrow +\infty} F^{uu}(y_1, y_0, x|w) \right] \\ &= \gamma [\varepsilon p_{x|w} + (1 - \varepsilon) p_{x|w}] + (1 - \gamma) [\varepsilon p_{x|w} + (1 - \varepsilon) p_{x|w}] \\ &= p_{x|w}, \end{aligned}$$

where the third equality follows by the condition 2 in the proof of Lemma 11.

Verifying Condition 3: Similar to the proof of condition 2, condition 3 follows by the fact that all the cdfs F^{ul} , F^{lu} , F^{uu} , and F^{ll} satisfy condition 3 as argued in Lemma 11. Hence it follows that their convex combination $\tilde{\mathbb{P}}$ also satisfies this condition.

Verifying Condition 4: As in the proof of Lemma 11, we established propensity score functions p^{ul} , p^{lu} , p^{uu} , and p^{ll} under the cdfs F^{ul} , F^{lu} , F^{uu} , and F^{ll} , respectively. Applying Lemma 12 to the conditional mixture distribution $\tilde{\mathbb{P}}$ gives the propensity score as below

$$\begin{aligned} & \tilde{\mathbb{E}}(X \mid Y_1, Y_0, W = w) \\ &= \omega^{ul}(Y_1, Y_0)p^{ul}(Y_1, Y_0, w) + \omega^{lu}(Y_1, Y_0)p^{lu}(Y_1, Y_0, w) + \omega^{uu}(Y_1, Y_0)p^{uu}(Y_1, Y_0, w) + \omega^{ll}(Y_1, Y_0)p^{ll}(Y_1, Y_0, w) \end{aligned}$$

almost surely under $\tilde{\mathbb{P}}$, where $\omega^k(Y_1, Y_0) \in [0, 1]$, and $\sum_k \omega^k(Y_1, Y_0) = 1$ almost surely under $\tilde{\mathbb{P}}$ for $k \in \{ul, lu, uu, ll\}$. Since we have argued that $p^k(u, v, w) \in [\underline{c}, \bar{c}]$ for $(u, v) \in \mathbb{R}^2$, therefore, $\tilde{\mathbb{E}}(X \mid Y_1, Y_0, W = w) \in [\underline{c}, \bar{c}]$ almost surely under $\tilde{\mathbb{P}}$, which concludes the proof. \square

C.3 Proofs for Section 4.2

Proof of Theorem 3. First, we prove this proposition when Assumption 2 holds. Fix an arbitrary $(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)$. By the definition of $\mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)$, there exists a joint cdf $F_{Y_1, Y_0|X, W}$ that generates (F_1, F_0, C) and satisfies Assumption 2. By Lemma 2 and the monotonicity assumption, we have

$$\theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W}) \leq \theta(F_1, F_0, F_{Y,X,W}) \leq \theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W}).$$

Since (F_1, F_0, C) is arbitrary, we have

$$\theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W}) \leq \inf_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W})$$

and

$$\sup_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) \leq \theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W}).$$

Furthermore, as demonstrated by Theorem 1, $(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, C)$ and $(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, C)$ are contained in the identified set for any copula $C \in \mathcal{C}_{1,0|X,W}$. This implies

$$\inf_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) \leq \theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W})$$

and

$$\sup_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) \geq \theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W}).$$

Thus we conclude that

$$\inf_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) = \theta(\bar{F}_{Y_1|W}, \underline{F}_{Y_0|W}, F_{Y,X,W})$$

$$\sup_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta(F_1, F_0, F_{Y,X,W}) = \theta(\underline{F}_{Y_1|W}, \bar{F}_{Y_0|W}, F_{Y,X,W}).$$

Note that Theorem 1 also implies that $(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \bar{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \bar{F}_{Y_0|W}, C_{1,0|X,W})$ belongs to the identified set $\mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)$ for each $(\varepsilon, \gamma) \in [0, 1]^2$. By the continuity of the mapping $(\varepsilon, \gamma) \mapsto \theta(\varepsilon \underline{F}_{Y_1|W} + (1 - \varepsilon) \bar{F}_{Y_1|W}, \gamma \underline{F}_{Y_0|W} + (1 - \gamma) \bar{F}_{Y_0|W})$ and the definition of $\mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)$, the sharpness of the interior then follows by the intermediate value theorem. The proof follows the same arguments when imposing Assumption 3, where we use Theorem 2 instead of Theorem 1. \square

Proof of Lemma 3. Part 1: Because $Q_{Y_x}(U) \sim Y_x$ when $U \sim \text{Unif}(0, 1)$, we can write

$$\mathbb{E}[Y_x] = \mathbb{E}[Q_{Y_x}(U)] = \int_0^1 Q_{Y_x}(\tau) d\tau = \int_0^1 \theta_Q(F_{Y_x}; \tau) d\tau.$$

By the proof of Part 2 below, $\theta_Q(F_{Y_x}; \tau)$ is increasing in F_{Y_x} for all $\tau \in (0, 1)$. This implies $\theta_{\mathbb{E}}(F_{Y_x}) =$

$\int_0^1 \theta_Q(F_{Y_x}; \tau) d\tau$ is also increasing. Continuity follows from

$$\begin{aligned}\theta_{\mathbb{E}}(\varepsilon \underline{F}_{Y_x} + (1 - \varepsilon) \overline{F}_{Y_x}) &= \int y d(\varepsilon \underline{F}_{Y_x}(y) + (1 - \varepsilon) \overline{F}_{Y_x}(y)) \\ &= \int y d(\varepsilon \underline{F}_{Y_x}(y)) + \int y d((1 - \varepsilon) \overline{F}_{Y_x}(y)) \\ &= \varepsilon \theta_{\mathbb{E}}(\underline{F}_{Y_x}) + (1 - \varepsilon) \theta_{\mathbb{E}}(\overline{F}_{Y_x})\end{aligned}$$

being continuous in ε over $\varepsilon \in [0, 1]$.

Part 2: Suppose $F_{Y_x}(y) \leq F'_{Y_x}(y)$ for all $y \in \mathbb{R}$. Therefore, for any $\tau \in (0, 1)$, $\{y \in \mathbb{R} : F_{Y_x}(y) \geq \tau\} \subseteq \{y \in \mathbb{R} : F'_{Y_x}(y) \geq \tau\}$. Hence,

$$\theta_Q(F_{Y_x}; \tau) = \inf\{y \in \mathbb{R} : F_{Y_x}(y) \geq \tau\} \geq \inf\{y \in \mathbb{R} : F'_{Y_x}(y) \geq \tau\} = \theta_Q(F'_{Y_x}; \tau).$$

Since $F_{Y_x} \succeq F'_{Y_x}$, we have that $\theta_Q(F_{Y_x}; \tau)$ is increasing in F_{Y_x} .

Part 3: Suppose $F_{Y_x}(y) \leq F'_{Y_x}(y)$ for all $y \in \mathbb{R}$. Denote by

$$\begin{aligned}F_{Y_x|X}(y | 1 - x) &= \frac{F_{Y_x}(y) - F_{Y|X}(Y | X)p_x}{p_{1-x}}, \\ F'_{Y_x|X}(y | 1 - x) &= \frac{F'_{Y_x}(y) - F_{Y|X}(Y | X)p_x}{p_{1-x}}.\end{aligned}$$

Then $F_{Y_x|X}(\cdot | 1 - x) \succeq F'_{Y_x|X}(\cdot | 1 - x)$ and, by Part 2,

$$\theta_{CQ}(F_{Y_x}; \tau) = \theta_Q(F_{Y_x|X}(\cdot | 1 - x); \tau) \geq \theta_Q(F'_{Y_x|X}(\cdot | 1 - x); \tau) = \theta_{CQ}(F'_{Y_x}; \tau)$$

for any $\tau \in (0, 1)$. Therefore, θ_{CQ} is increasing in F_{Y_x} .

Part 4: Suppose $F_{Y_1|W} \succeq F_{Y'_1|W}$. This implies

$$\int y dF_{Y_1|W}(y | w) \geq \int y dF_{Y'_1|W}(y | w) \text{ for all } w \in \text{supp}(W)$$

which in turn implies

$$\mathbb{1}\left(\int y dF_{Y_1|W}(y | w) - \mathbb{E}[Y_0 | W = w] \leq z\right) \leq \mathbb{1}\left(\int y dF_{Y'_1|W}(y | w) - \mathbb{E}[Y_0 | W = w] \leq z\right)$$

for all $w \in \text{supp}(W)$ and hence

$$\mathbb{P}\left(\int y dF_{Y_1|W}(y | W) - \mathbb{E}[Y_0 | W] \leq z\right) \leq \mathbb{P}\left(\int y dF_{Y'_1|W}(y | W) - \mathbb{E}[Y_0 | W] \leq z\right).$$

The first statement holds by Part 1 of this lemma, the second holds directly, and the third by integrating over the distribution of W . Therefore, this last cdf is decreasing in $F_{Y_1|W}$. By Part 2 of this lemma, its corresponding quantile will be decreasing in $F_{Y_1|W}$. This parameter is decreasing in $F_{Y_0|W}$ because of the minus sign inside the CATE.

Part 5: Suppose $F_{Y_x} \succeq F'_{Y_x}$ for $x \in \{0, 1\}$. Then, $F_{Y_x}(y) \leq F'_{Y_x}(y)$ for all $y \in \mathbb{R}$. Therefore, for any $(y_1, y_0) \in \mathbb{R}^2$ and copula C

$$\theta(F_{Y_1}, F_{Y_0}, C; y_1, y_0) = C(F_{Y_1}(y_0), F_{Y_0}(y_0)) \leq C(F'_{Y_1}(y_0), F'_{Y_0}(y_0)) = \theta(F'_{Y_1}, F'_{Y_0}, C; y_1, y_0)$$

because C , as a copula, is nondecreasing in its arguments. We conclude that this parameter is decreasing in both F_{Y_1} and F_{Y_0} .

Part 6: We begin by showing that $(Y_1, Y_0) \sim (Q_{Y_1}(U_1), Q_{Y_0}(U_0))$ where (U_1, U_0) have joint cdf C . To

see this, note that $F_{Y_1, Y_0}(y_1, y_0) = C(F_{Y_1}(y_1), F_{Y_0}(y_0))$ by Sklar's Theorem. Also,

$$\begin{aligned} F_{Y_1, Y_0}(y_1, y_0) &= C(F_{Y_1}(y_1), F_{Y_0}(y_0)) \\ &= \mathbb{P}(U_1 \leq F_{Y_1}(y_1), U_0 \leq F_{Y_0}(y_0)) \\ &= \mathbb{P}(Q_{Y_1}(U_1) \leq y_1, Q_{Y_0}(U_0) \leq y_0), \end{aligned}$$

where the third equality follows from Lemma 6.1.

Based on this, we can write the functional as

$$\begin{aligned} \theta_{\text{DTE}}(F_{Y_1}, F_{Y_0}, C; z) &= \mathbb{E}[\mathbb{1}(Y_1 - Y_0 \leq z)] \\ &= \mathbb{E}[\mathbb{1}(Q_{Y_1}(U_1) - Q_{Y_0}(U_0) \leq z)] \\ &= \int \mathbb{1}(Q_{Y_1}(u_1) - Q_{Y_0}(u_0) \leq z) dC(u_1, u_0). \end{aligned}$$

Now suppose that $F_{Y_1} \succeq F'_{Y_1}$. By Part 2 above, this implies that $Q_{Y_1}(u_1) \geq Q'_{Y_1}(u_1)$ for all $u_1 \in (0, 1)$ and thus

$$\begin{aligned} \theta_{\text{DTE}}(F_{Y_1}, F_{Y_0}, C; z) &= \int \mathbb{1}(Q_{Y_1}(u_1) - Q_{Y_0}(u_0) \leq z) dC(u_1, u_0) \\ &\leq \int \mathbb{1}(Q'_{Y_1}(u_1) - Q_{Y_0}(u_0) \leq z) dC(u_1, u_0) \\ &= \theta_{\text{DTE}}(F'_{Y_1}, F_{Y_0}, C; z). \end{aligned}$$

Therefore, $\theta_{\text{DTE}}(F_{Y_1}, F_{Y_0}, C; z)$ is decreasing in F_{Y_1} . An analogous argument shows that it is increasing in F_{Y_0} . \square

D Proofs for Section 5

Proof of Proposition 2. By Lemma 2 and the monotonicity of copulas in their arguments, we have that

$$\begin{aligned} &\sup_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y, X, W}; c)} \theta_{\text{CDF}}(F_1, F_0, C, F_{Y, X, W}; y_1, y_0) \\ &\leq \sup_{C \in \mathcal{C}_{1, 0|X, W}} \theta_{\text{CDF}}(\overline{F}_{Y_1|X, W}, \overline{F}_{Y_0|X, W}, C, F_{Y, X, W}; y_1, y_0) \\ &= \theta_{\text{CDF}}(\overline{F}_{Y_1|X, W}, \overline{F}_{Y_0|X, W}, \overline{C}, F_{Y, X, W}; y_1, y_0). \end{aligned}$$

The equality follows from the Fréchet-Hoeffding bounds. Similarly,

$$\begin{aligned} &\inf_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y, X, W}; c)} \theta_{\text{CDF}}(F_1, F_0, C, F_{Y, X, W}; y_1, y_0) \\ &\geq \inf_{C \in \mathcal{C}_{1, 0|X, W}} \theta_{\text{CDF}}(\underline{F}_{Y_1|X, W}, \underline{F}_{Y_0|X, W}, C, F_{Y, X, W}; y_1, y_0) \\ &= \theta_{\text{CDF}}(\underline{F}_{Y_1|X, W}, \underline{F}_{Y_0|X, W}, \underline{C}, F_{Y, X, W}; y_1, y_0). \end{aligned}$$

To show sharpness of the interval (7), consider the following choices for conditional cdfs and copulas:

$$(\varepsilon_1 \underline{F}_{Y_1|X, W} + (1 - \varepsilon_1) \overline{F}_{Y_1|X, W}, \varepsilon_2 \underline{F}_{Y_0|X, W} + (1 - \varepsilon_2) \overline{F}_{Y_0|X, W}, \varepsilon_3 \underline{C} + (1 - \varepsilon_3) \overline{C})$$

for $\varepsilon := (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in [0, 1]^3$. For any $\varepsilon \in [0, 1]^3$, this triple belongs to $\mathcal{I}_0^{\text{marg}}(F_{Y, X, W}; c)$. Setting $\varepsilon = (0, 0, 0)$ and $\varepsilon = (1, 1, 1)$ yields the upper and lower bounds of the interval, so the bounds are sharp. To show the

interior is sharp, consider the function

$$\begin{aligned} \varepsilon &\mapsto \theta_{\text{CDF}} \left(\begin{array}{c} \varepsilon_1 \underline{F}_{Y_1|X,W} + (1 - \varepsilon_1) \overline{F}_{Y_1|X,W}, \\ \varepsilon_2 \underline{F}_{Y_0|X,W} + (1 - \varepsilon_2) \overline{F}_{Y_0|X,W}, \\ \varepsilon_3 \underline{C} + (1 - \varepsilon_3) \overline{C}, \\ F_{Y,X,W} \end{array} ; y_1, y_0 \right) \\ &= \varepsilon_3 \mathbb{E} \left[\max \left\{ \begin{array}{l} \varepsilon_1 \underline{F}_{Y_1|X,W}(y_1 | X, W) + (1 - \varepsilon_1) \overline{F}_{Y_1|X,W}(y_1 | X, W) \\ + \varepsilon_2 \underline{F}_{Y_0|X,W}(y_0 | X, W) + (1 - \varepsilon_2) \overline{F}_{Y_0|X,W}(y_0 | X, W) - 1, 0 \end{array} \right\} \right] \\ &\quad + (1 - \varepsilon_3) \mathbb{E} \left[\min \left\{ \begin{array}{l} \varepsilon_1 \underline{F}_{Y_1|X,W}(y_1 | X, W) + (1 - \varepsilon_1) \overline{F}_{Y_1|X,W}(y_1 | X, W), \\ \varepsilon_2 \underline{F}_{Y_0|X,W}(y_0 | X, W) + (1 - \varepsilon_2) \overline{F}_{Y_0|X,W}(y_0 | X, W) \end{array} \right\} \right]. \end{aligned}$$

This mapping is continuous in ε_3 . It is also continuous in ε_1 and ε_2 since the functions $(u, v) \mapsto \underline{C}(u, v)$ and $(u, v) \mapsto \overline{C}(u, v)$ are both continuous, and by the dominated convergence theorem. Therefore, by the intermediate value theorem, all values in the interval (7) are attained and thus the identified set is this interval. \square

Proof of Proposition 3. By lemmas 2 and 3.5

$$\begin{aligned} &\sup_{(F_1, F_0, C) \in \mathcal{I}_0^{\text{marg}}(F_{Y,X,W}; c)} \theta_{\text{DTE}}(F_{Y_1|X,W}, F_{Y_0|X,W}, C_{1,0|X,W}, F_{Y,X,W}; z) \\ &\leq \sup_{C \in \mathcal{C}_{1,0|X,W}} \theta_{\text{DTE}}(\overline{F}_{Y_1|X,W}, \underline{F}_{Y_0|X,W}, C, F_{Y,X,W}; z). \end{aligned}$$

By Lemma 2.1 in Fan and Park (2010),

$$\begin{aligned} &\sup_{C \in \mathcal{C}_{1,0|X,W}} \theta_{\text{DTE}}(\overline{F}_{Y_1|X,W}, \underline{F}_{Y_0|X,W}, C, F_{Y,X,W}; z) \\ &\leq 1 + \mathbb{E} \left[\min \left\{ \inf_{y \in \mathbb{R}} \left(\overline{F}_{Y_1|X,W}(y | X, W) - \underline{F}_{Y_0|X,W}(y - z | X, W) \right), 0 \right\} \right]. \end{aligned} \tag{51}$$

This bound can be attained because the cdf pair $(\overline{F}_{Y_1|X,W}, \underline{F}_{Y_0|X,W})$ is attainable by Theorem 1, and the bound in (51) is attained by Lemma 2.1 in Fan and Park (2010) since the set of conditional copulas under marginal c -dependence is unrestricted.

Similar proof can be used to show that the lower bound

$$\mathbb{E} \left[\max \left\{ \sup_{y \in \mathbb{R}} \left(\underline{F}_{Y_1|X,W}(y | X, W) - \overline{F}_{Y_0|X,W}(y - z | X, W) \right), 0 \right\} \right]$$

is sharp as well. \square

E Appendix: Explicit Bounds on Expected Potential Outcomes

Lemma 13. Let Y be random variable with cdf F and quantile function Q . Suppose $\mathbb{E}(|Y|) < \infty$. Then, for $a \in (0, 1)$:

$$\begin{aligned} \int_0^a Q(u) du &= \mathbb{E}[Y | Y \leq Q(a)]F(Q(a)) - Q(a)(F(Q(a)) - a) \\ \int_a^1 Q(u) du &= \mathbb{E}[Y | Y > Q(a)](1 - F(Q(a))) + Q(a)(F(Q(a)) - a). \end{aligned}$$

If $\mathbb{P}(Y = Q(a)) = 0$, then

$$\begin{aligned}\int_0^a Q(u) du &= \mathbb{E}[Y \mid Y \leq Q(a)]a \\ \int_a^1 Q(u) du &= \mathbb{E}[Y \mid Y > Q(a)](1-a).\end{aligned}$$

Lemma 14. Let Assumption 1 hold. Then,

$$\begin{aligned}& \int y d\bar{F}_{Y_1|W}(y \mid w) \\ &= \frac{p_{1|w}}{\underline{c}} \left[\mathbb{E}[Y \mid Y \leq \bar{Q}_1, X = 1, W = w] F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{Q}_1 (F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1) \right] \\ &+ \frac{p_{1|w}}{\bar{c}} \left[\mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w] (1 - F_{Y|X,W}(\bar{Q}_1 \mid 1, w)) + \bar{Q}_1 (\bar{F}_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1) \right],\end{aligned}\tag{52}$$

$$\begin{aligned}& \int y d\underline{F}_{Y_1|W}(y \mid w) \\ &= \frac{p_{1|w}}{\bar{c}} \left[\mathbb{E}[Y \mid Y \leq \underline{Q}_1, X = 1, W = w] F_{Y|X,W}(\underline{Q}_1 \mid 1, w) - \underline{Q}_1 (F_{Y|X,W}(\underline{Q}_1 \mid 1, w) - \underline{\tau}_1) \right] \\ &+ \frac{p_{1|w}}{\underline{c}} \left[\mathbb{E}[Y \mid Y > \underline{Q}_1, X = 1, W = w] (1 - F_{Y|X,W}(\underline{Q}_1 \mid 1, w)) + \underline{Q}_1 (F_{Y|X,W}(\underline{Q}_1 \mid 1, w) - \underline{\tau}_1) \right]\end{aligned}\tag{53}$$

and

$$\begin{aligned}& \int y d\bar{F}_{Y_0|W}(y \mid w) \\ &= \frac{p_{0|w}}{1-\bar{c}} \left[\mathbb{E}[Y \mid Y \leq \bar{Q}_0, X = 0, W = w] F_{Y|X,W}(\bar{Q}_0|0, w) - \bar{Q}_0 (F_{Y|X,W}(\bar{Q}_0|0, w) - \bar{\tau}_0) \right] \\ &+ \frac{p_{0|w}}{1-\underline{c}} \left[\mathbb{E}[Y \mid Y > \bar{Q}_0, X = 0, W = w] (1 - F_{Y|X,W}(\bar{Q}_0|0, w)) + \bar{Q}_0 (F_{Y|X,W}(\bar{Q}_0|0, w) - \bar{\tau}_0) \right],\end{aligned}\tag{54}$$

$$\begin{aligned}& \int y d\underline{F}_{Y_0|W}(y \mid w) \\ &= \frac{p_{0|w}}{1-\underline{c}} \left[\mathbb{E}[Y \mid Y \leq \underline{Q}_0, X = 0, W = w] F_{Y|X,W}(\underline{Q}_0|0, w) - \underline{Q}_0 (F_{Y|X,W}(\underline{Q}_0|0, w) - \underline{\tau}_0) \right] \\ &+ \frac{p_{0|w}}{1-\bar{c}} \left[\mathbb{E}[Y \mid Y > \underline{Q}_0, X = 0, W = w] (1 - F_{Y|X,W}(\underline{Q}_0|0, w)) + \underline{Q}_0 (F_{Y|X,W}(\underline{Q}_0|0, w) - \underline{\tau}_0) \right].\end{aligned}\tag{55}$$

If Y is continuously distributed conditionally on (X, W) , then these expressions simplify to

$$\begin{aligned}\int y d\bar{F}_{Y_1|W}(y \mid w) &= \mathbb{E}[Y \mid Y \leq \bar{Q}_1, X = 1, W = w] \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \\ &+ \mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w] \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} \\ \int y d\underline{F}_{Y_1|W}(y \mid w) &= \mathbb{E}[Y \mid Y \leq \underline{Q}_1, X = 1, W = w] \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} \\ &+ \mathbb{E}[Y \mid Y > \underline{Q}_1, X = 1, W = w] \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}}\end{aligned}$$

and

$$\begin{aligned}
\int y d\bar{F}_{Y_0|W}(y | w) &= \mathbb{E}[Y | Y \leq \bar{Q}_0, X = 0, W = w] \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}} \\
&\quad + \mathbb{E}[Y | Y > \bar{Q}_0, X = 0, W = w] \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \\
\int y d\underline{F}_{Y_0|W}(y | w) &= \mathbb{E}[Y | Y \leq \underline{Q}_0, X = 0, W = w] \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \\
&\quad + \mathbb{E}[Y | Y > \underline{Q}_0, X = 0, W = w] \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}}
\end{aligned}$$

where \bar{Q}_x , \underline{Q}_x , $\bar{\tau}_x$, and $\underline{\tau}_x$ for $x = 0, 1$ are defined in Appendix A.

E.1 Proofs of Analytical Bounds on Expectations

Proof of Lemma 13. First we consider the equality involving $\int_0^a Q(u) du$. Note that

$$\begin{aligned}
\int_0^a Q(u) du &= \int_0^1 Q(u) \mathbb{1}[u \leq a] du \\
&= \int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a), u \leq a] du \\
&= \int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a)] du - \int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a), u > a] du \\
&= \int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a)] du - \int_0^1 Q(u) \mathbb{1}[Q(u) = Q(a), u > a] du \\
&= \int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a)] du - Q(a) \int_0^1 \mathbb{1}[Q(u) \leq Q(a), u > a] du,
\end{aligned}$$

where the second, the fourth, and the last line follow by the monotonicity of quantile function $Q(\cdot)$.

The first term in the last line can be written as below:

$$\int_0^1 Q(u) \mathbb{1}[Q(u) \leq Q(a)] du = \mathbb{E}(Y \mathbb{1}[Y \leq Q(a)]) = \mathbb{E}(Y | Y \leq Q(a)) F(Q(a)),$$

where the first equality follows by that $Q(U)$ has the same distribution as Y if U is uniformly distributed over $[0, 1]$. To expand the second term, note that $\{u : Q(u) \leq Q(a), u > a\}$ is a half-open interval with the left endpoint a excluded, and right endpoint $\sup\{u : Q(u) \leq Q(a)\}$ included in the interval due to the left-continuity of quantile function $Q(\cdot)$. So we have

$$\begin{aligned}
Q(a) \int_0^1 \mathbb{1}[Q(u) \leq Q(a), u > a] du &= Q(a)(\sup\{u : Q(u) \leq Q(a)\} - a) \\
&= Q(a)(\sup\{u : u \leq F(Q(a))\} - a) \\
&= Q(a)(F(Q(a)) - a),
\end{aligned}$$

where the second line holds by Lemma 6.1. Given the above derivations, we conclude that

$$\int_0^a Q(u) du = \mathbb{E}[Y | Y \leq Q(a)] F(Q(a)) - Q(a)(F(Q(a)) - a),$$

as desired.

Regarding the second equality involving $\int_a^1 Q(u) du$, note that $\int_0^1 Q(u) du = \mathbb{E}[Q(U)] = \mathbb{E}[Y]$. This

implies

$$\begin{aligned}
\int_a^1 Q(u) du &= \int_0^1 Q(u) du - \int_0^a Q(u) du \\
&= \mathbb{E}[Y] - \mathbb{E}[Y | Y \leq Q(a)]F(Q(a)) + Q(a)(F(Q(a)) - a) \\
&= \mathbb{E}[Y | Y > Q(a)](1 - F(Q(a))) + Q(a)(F(Q(a)) - a),
\end{aligned}$$

where the last line follows by the law of iterated expectation. So the second equality is established.

When $\mathbb{P}(Y = Q(a)) = 0$, the the CDF F is continuous at $Q(a)$, which implies that $F(Q(a)) = a$ by Lemma 6.2. Therefore,

$$\int_0^a Q(u) du = \mathbb{E}[Y | Y \leq Q(a)]F(Q(a)) - Q(a)(F(Q(a)) - a) = \mathbb{E}[Y | Y \leq Q(a)]a,$$

and similar arguments can be applied to $\int_a^1 Q(u) du$ as well. Therefore we have established the desired result. \square

Proof of Lemma 14. We prove the claim for $\int_0^1 y d\bar{F}_{Y_1|W}(y | w)$, and note that the claims for the other terms can be derived analogously.

Let $U \sim \text{Unif}(0, 1)$, then $Q_{Y_1|W}(U | 1, w)$ has the distribution $\bar{F}_{Y_1|W}(\cdot | 1, w)$, which implies

$$\begin{aligned}
\int y d\bar{F}_{Y_1|W}(y | w) &= \int_0^1 Q_{Y_1|W}(\tau | w) d\tau \\
&= \int_0^1 Q_{Y|X,W} \left(\frac{\underline{c}\tau}{p_{1|w}} | 1, w \right) \mathbb{1} \left[\tau \leq \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \right] d\tau \tag{56}
\end{aligned}$$

$$+ \int_0^1 Q_{Y|X,W} \left(\frac{p_{1|w} - \bar{c} + \bar{c}\tau}{p_{1|w}} | 1, w \right) \mathbb{1} \left[\tau > \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \right] d\tau \tag{57}$$

We expand the term (56) below:

$$\begin{aligned}
&\int_0^1 Q_{Y|X,W} \left(\frac{\underline{c}\tau}{p_{1|w}} | 1, w \right) \mathbb{1} \left[\tau \leq \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \right] d\tau \\
&= \frac{p_{1|w}}{\underline{c}} \int_0^{\bar{\tau}_1} Q_{Y|X,W}(u | 1, w) du \\
&= \frac{p_{1|w}}{\underline{c}} [\mathbb{E}[Y | Y \leq Q_{Y|X,W}(\bar{\tau}_1 | 1, w), X = 1, W = w] F_{Y|X,W}(Q_{Y|X,W}(\bar{\tau}_1 | 1, w) | 1, w)] \\
&\quad - \frac{p_{1|w}}{\underline{c}} Q_{Y|X,W}(\bar{\tau}_1 | 1, w) [F_{Y|X,W}(Q_{Y|X,W}(\bar{\tau}_1 | 1, w) | 1, w) - \bar{\tau}_1] \\
&= \frac{p_{1|w}}{\underline{c}} [\mathbb{E}[Y | Y \leq \bar{Q}_1, X = 1, W = w] F_{Y|X,W}(\bar{Q}_1 | 1, w) - \bar{Q}_1 (F_{Y|X,W}(\bar{Q}_1 | 1, w) - \bar{\tau}_1)]. \tag{58}
\end{aligned}$$

The first equality uses the changes of variable $u = \underline{c}\tau/p_{1|w}$ and recall that

$$\bar{\tau}_1 = \frac{\underline{c}}{p_{1|w}} \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}}.$$

The second equality follows by Lemma 13, and the last line holds by recalling that $\bar{Q}_1 = Q_{Y|X,W}(\bar{\tau}_1 | 1, w)$.

Similarly, we can expand the term (57) below:

$$\begin{aligned}
& \int_0^1 Q_{Y|X,W} \left(\frac{p_{1|w} - \bar{c} + \bar{c}\tau}{p_{1|w}} \mid 1, w \right) \mathbb{1} \left[\tau > \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} \right] d\tau \\
&= \int_{\bar{\tau}_1}^1 Q_{Y|X,W} (u \mid 1, w) du \\
&= \frac{p_{1|w}}{\bar{c}} \left[\mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w](1 - F_{Y|X,W}(\bar{Q}_1 \mid 1, w)) + \bar{Q}_1 (F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1) \right], \quad (59)
\end{aligned}$$

where we use the change of variable $u = 1 - \frac{(1-\tau)\bar{c}}{p_{1|w}}$ in the second line.

Combining the above results, we can combine (58) and (59) to obtain the analytical formula of $\int y d\bar{F}_{Y_1|W}$:

$$\begin{aligned}
& \int y d\bar{F}_{Y_1|W}(y \mid w) \\
&= \frac{p_{1|w}}{\underline{c}} \left[\mathbb{E}[Y \mid Y \leq \bar{Q}_1, X = 1, W = w] F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{Q}_1 (F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1) \right] \\
&\quad + \frac{p_{1|w}}{\bar{c}} \left[\mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w](1 - F_{Y|X,W}(\bar{Q}_1 \mid 1, w)) + \bar{Q}_1 (F_{Y|X,W}(\bar{Q}_1 \mid 1, w) - \bar{\tau}_1) \right].
\end{aligned}$$

Finally, we note that if Y is continuously distributed conditional on (X, W) , then

$$F_{Y|X,W}(\bar{Q}_1 \mid 1, w) = \bar{\tau}_1,$$

which implies

$$\begin{aligned}
\int y d\bar{F}_{Y_1|W}(y \mid w) &= \frac{p_{1|w}}{\underline{c}} \mathbb{E}[Y \mid Y \leq \bar{Q}_1, X = 1, W = w] \bar{\tau}_1 + \frac{p_{1|w}}{\bar{c}} \mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w](1 - \bar{\tau}_1) \\
&= \mathbb{E}[Y \mid Y \leq \bar{Q}_1, X = 1, W = w] \frac{\bar{c} - p_{1|w}}{\bar{c} - \underline{c}} + \mathbb{E}[Y \mid Y > \bar{Q}_1, X = 1, W = w] \frac{p_{1|w} - \underline{c}}{\bar{c} - \underline{c}},
\end{aligned}$$

as desired. \square