Sharp Bounds and Inference in Sample Selection Models with Treatment Endogeneity*

Yingying Dong[†] Phillip Heiler[‡]

This version: June 9, 2026

This paper provides partial identification and inference for treatment effects in nonparametric sample selection models with endogenous treatment and (weak) sample selection monotonicity. Outcomes are observed only for a non-randomly selected subsample and treatment is endogenous because of noncompliance with assignment. The proposed bounds for intensive margin treatment effects among compliers are sharp and tighter than those of Chen and Flores (2015). For inference, we develop semiparametrically efficient orthogonal moments and a debiased machine learning procedure that permits valid root-n inference under high-dimensional covariates and/or flexible functional forms. Simulation results indicate good finite sample performance. Applications to Job Corps and the Oregon Health Insurance Experiment show that the method can deliver substantially tighter effect bounds and confidence intervals than existing alternatives.

Keywords: Debiased/double machine learning; Lee bounds; Partial identification; Principal strata; Noncompliance; Sample selection

JEL classification: C13, C14, C21

*We would like to thank Otavio Bartalotti, Giovanni Mellace, Vitor Possebom, Vira Semenova, and the participants at the IAAE 2025 and USC-UCLA Mini Workshop for helpful comments and discussion. Phillip Heiler would also like to thank the Danish National Research Foundation (DNRF186) and the Independent Research Fund Denmark for support (DFF 3099-00077B). All remaining errors are ours.

[†]University of California Irvine, Department of Economics, Irvine, CA, 92697, USA. email: yyd@uci.edu

[‡]Aarhus University. Department of Economics and Business Economics, Aarhus Center for Econometrics (ACE), TrygFonden's Centre for Child Research, Universitetsbyen 51, 8000 Aarhus C, Denmark. email: pheiler@econ.au.dk

1. Introduction

This paper provides sharp bounds and inference for the causal effect of a binary treatment in settings with nonrandom sample selection and endogenous treatment take-up – two prominent features of many empirical applications. Our paper can be viewed from two complementary perspectives: From the perspective of the sample selection literature, we extend intensive margin bounds from reduced-form effects of treatment assignment to complier treatment effects in the presence of treatment take-up endogeneity. From the perspective of the instrumental variable/local average treatment effect (IV/LATE) literature, we extend the standard LATE framework to allow for sample selection, or partial observability of the outcome. Thus, this paper targets the intensive margin analogue of LATE for always-selected compliers. The framework nests the widely used Lee (2009) bounds as a special case. Identification is obtained under standard IV/LATE assumptions together with weak sample selection monotonicity for treatment compliers. We also develop semiparametrically efficient orthogonal moments and debiased machine learning (DML) estimators that allow valid root-n inference on the treatment effect and its sharp bounds under high-dimensional covariates and flexible functional forms.

Partial identification of causal effects under sample selection or partial observability has been extensively studied. Much of the existing work, including Lee (2009), derives bounds for the effect of a randomly assigned instrument without accounting for endogenous treatment take-up, and is therefore limited to intention-to-treat (ITT) effects when assignment and treatment receipt differ, see, e.g., Horowitz and Manski (2000), Zhang and Rubin (2003), Imai (2008), Huber and Mellace (2015), Heiler (2024), Heiler et al. (2024), Sun et al. (2024), or Lee and Liu (2025).[1]

[1] Horowitz and Manski (2000) propose nonparametric bounds for randomized experiments with missing covariate and outcome data. Zhang and Rubin (2003) develop bounds for the survivor average causal effect, targeting individuals who are always selected regardless of treatment status using principal stratification (Frangakis and Rubin, 2002). They provide both assumption-free bounds and bounds under sample selection monotonicity and stochastic dominance assumptions. Their monotonicity based bounds are now regularly referred to as "Lee bounds" or "Zhang-Rubin-Lee bounds" (Andersen et al., 2023). Imai (2008) proves the sharpness of the Zhang and Rubin (2003) bounds and extends them to quantile treatment effects. Huber and Mellace (2015) derive bounds for additional subpopulations, such as individuals selected only under treatment or only under control, and for the observed subpopulation. Heiler (2024) provides heterogeneous treatment effect bounds. Heiler et al. (2024) provide identification for principal strata and extensive and intensive margin bounds. Sun et al. (2024) consider bounds for type-specific potential outcome means, assuming exogenous treatment assignment and using an excluded variable that shifts selection but not potential outcomes. Lee and Liu (2025) consider intensive margin bounds for continuous treatments.

In practice, however, realized treatment take-up is often endogenous in both experimental and quasi-experimental settings as assignment, eligibility, or encouragement may shift participation without perfectly determining treatment. For instance, in our empirical applications a nontrivial fraction of individuals do not comply with the assignment. In the National Job Corps (JC) Study, 26.2% of the individuals provided with access to training did not enroll in JC, while 4.4% of individuals assigned to the control group eventually enrolled in JC after randomization (Schochet et al., 2008). In the Oregon Health Insurance Experiment (OHIE), only 30% of those eligible to apply for Medicaid successfully enrolled, and a small share of controls also enrolled (Finkelstein et al., 2012). Noncompliance of these magnitudes can substantially affect treatment effect bounds.

Without sample selection, ITT and complier treatment effect differ only by a scaling factor, the share of compliers (Angrist et al., 1996). We show that this simple relationship breaks down under sample selection: Sharp treatment effect bounds cannot be obtained by merely scaling sharp bounds of ITT effects via primitive probabilities. Our analysis therefore fills an important gap by delivering identification and inference for sharp treatment effect bounds in the more realistic case where both endogenous sample selection and noncompliance (implying treatment endogeneity) are present.

Other papers have leveraged IV/LATE-type assumptions for partial identification of various parameters in sample selection models: Lechner and Melly (2010) derive bounds for mean and quantile treatment effects for observable subpopulations, such as the "treated and selected". Christelis and Messina (2019) provide bounds on potential outcome distributions under selected samples and a monotone IV assumption (Manski, 1997).

Within this literature, the closest contribution to ours is Chen and Flores (2015; CF). To our knowledge, CF is the only paper studying bounds in a setting with both sample selection and noncompliance that also provides statistical inference. Unlike the CF bounds, our bounds are sharp. Related identification contributions include Imai (2007)[2]

[2]The derivation of bounds in Imai (2007) is based on a principal-strata approach that is complementary to ours. However, the corresponding bound expressions imported from Imai (2008) appear to contain errors in the expressions for Q_0 and Q_1 in RESULT 1 on page 4. Hence, the resulting bounds cannot be reconciled with ours. Moreover, Imai (2007) does not develop estimation or inference, and sharpness is not formally established for the noncompliance-and-selection parameter. Sharpness is only indirectly suggested by the arguments in Imai (2008), which is formulated for ITT bounds under different assumptions. By contrast, we provide corrected expressions for the bounds, prove their sharpness, compare them with CF bounds, and develop a full estimation and inference framework that extends to settings with both discrete and continuous covariates and weak sample selection monotonicity.

and Bartalotti et al. (2023).[3] Relative to all three papers, we contribute along multiple dimensions: First, we allow for the inclusion of covariates, thereby accommodating unconfounded rather than independently assigned instruments, which is often more realistic in quasi-experimental settings. Even when the instrument is independently assigned, incorporating covariates can improve efficiency, much as in standard average treatment effect (ATE) estimation. Second, we relax their individual-level strong sample selection monotonicity to weak, i.e., covariate-dependent, monotonicity. This is empirically relevant: Semenova (2025) shows that strong sample selection monotonicity can be rejected in JC. We also find substantial violations in the OHIE. Third, we provide a simple covariate-profiling procedure for the targeted complier population at the intensive margin. This helps to assess both the external validity of the estimates and the substantive relevance of the subpopulation to which they apply, paralleling complier profiling approaches from the LATE literature (Abadie, 2003; Angrist, 2004; Singh and Sun, 2023). Fourth, we propose a semiparametric DML procedure that delivers valid root-n inference using generic nonparametric or machine learning methods for the nuisance components, allowing for potentially high-dimensional covariates and/or unknown functional forms. In strong contrast to Chen and Flores (2015), our estimators are asymptotically linear, and thus confidence intervals have the standard "estimate $\pm$ standard error $\times$ critical value" form and do not rely on alternative concepts such as half-median-unbiasedness (Chernozhukov et al., 2013).

Recent research has generalized estimation and inference of Lee-type bounds to high-dimensional or otherwise flexible settings using semiparametric DML methodology. Semenova (2025) extends Lee (2009) intensive margin bounds to multiple outcomes and provides orthogonal moments for debiased estimation. Heiler (2024) develops a DML procedure to obtain corresponding heterogeneous treatment effect bounds and inference under local misspecification. Heiler et al. (2024) provide refined DML methods for outer identification regions for intensive and extensive margin effects that admit regular infer-

[3] Bartalotti et al. (2023) study treatment-effect bounds under sample selection within a latent-index marginal treatment effect (MTE) framework. Their Proposition 5 derives sharp bounds for always-observed LATEs with multi-valued discrete instruments. In the binary-instrument, no-covariate case, their latent-index interval $p_0 < V \leq p_1$ coincides with our complier stratum as defined in Section 2. Under the same strong sample selection monotonicity, their Proposition 5 bounds are algebraically equivalent to our basic sharp bounds without covariates in Section 2. They informally discuss plug-in estimation for MTE bounds without covariates, but do not develop accompanying inference guarantees for the discrete-instrument always-observed LATE bounds in Proposition 5.

ence in a larger class of DGPs. These contributions all focus on inference for (conditional) ITT effects. We instead construct semiparametrically efficient orthogonal moment functions for the sharp bounds on the treatment effect for compliers at the intensive margin. They nest several leading moments from the literature, including for ITT effect bounds (Semenova 2025, under perfect compliance), LATE (Frölich 2007, in the absence of sample selection), and ATE (Hahn 1998, under perfect compliance and no sample selection).

A crucial difference to the literature on DML inference on Lee-type ITT bounds is that our sharp bounds require trimming based on a population whose quantiles are identified only *indirectly* via inversion of a compliance weighted cumulative distribution in the sense of Abadie (2003). The latter by itself is a linear, but not necessarily convex, combination of two conditional CDFs. As a result, debiasing relies on an implicit function characterization. We leverage this representation to impose learning rate requirements only on primary, reduced-form type conditional CDFs and conditional means instead of the counterfactual conditional quantiles used for trimming the relevant observed strata distributions. This also has the advantage of making it straightforward in practice to guarantee non-crossing properties for the conditional quantiles involved, e.g., via inversion of isotonic (distributional) regression (Henzi et al., 2021). Our implementation also follows this inversion strategy. Monte Carlo simulations suggest good coverage and power properties of the inference method in finite samples.

We apply and compare our method to evaluate the effects of Job Corps on earnings and Medicaid on healthcare utilization. Under the strong sample selection monotonicity assumption, our sharp bounds tighten the existing Chen and Flores (2015) bounds by 68.4% (JC) and 31.1% to 69.4% (OHIE, depending on outcome). For example, the sharp bounds without covariates for the intensive margin complier effect of JC on hourly wages are now $[0.019, 0.067]$ compared to $[-0.022, 0.130]$ for CF. Similar reductions apply to the 95%-confidence intervals for the effects. Their widths are reduced by 46.7% (JC) and 25.8% to 36.9% (OHIE, depending on outcome). Under weak sample selection monotonicity, our sharp DML bounds still tend to be contained by the restrictive CF bounds for most outcomes and, despite relying on weaker assumptions, continue to yield shorter confidence intervals of around 7.9% (JC) and 7.3% to 42.3% (OHIE, depending on outcome).

The rest of the paper is organized as follows: Section 2 derives the basic bounds without covariates and compares them with Lee and CF bounds. Section 3 extends these bounds to incorporate covariates. Section 4 develops estimation and inference. Section 5 presents the profiling method for the targeted complier population at the intensive margin. Section 6 provides the JC application. Section 7 concludes. The OHIE application, Monte Carlo simulations as well as proofs and supplementary derivations are in the Supplementary Appendix.

2. Basic Bounds Without Covariates

2.1. Model and Identification of Basic Bounds

Let $D \in \{0,1\}$ denote a binary treatment and $Z \in \{0,1\}$ a binary instrument. Let $S \in \{0,1\}$ be a sample selection indicator for a continuous outcome $Y \in \mathcal{Y} \subset \mathbb{R}$ that is observed only if $S = 1$. The observed data are independent draws of (SY, S, D, Z). For $d, z \in \{0,1\}$, let $Y_{d,z}$ and $S_{d,z}$ denote, respectively, the potential outcome and selection indicator that would be observed if treatment and the instrument were set exogenously to (d, z). Let D_z denote the potential treatment if Z were set to z. We impose the following IV assumptions:

Assumption 2.1 (IV)

2.1.1 *(Exclusion) For* $d = 0, 1$, $Y_{d,1} = Y_{d,0} =: Y_d$ *and* $S_{d,1} = S_{d,0} =: S_d$.

2.1.2 *(Independence)* $(Y_1, Y_0, S_1, S_0, D_1, D_0) \perp Z$.

2.1.3 *(Strong Treatment Response Monotonicity)* $P(D_1 \geq D_0) = 1$.

2.1.4 *(First Stage)* $P(D_1 \neq D_0) > 0$.

2.1.5 *(Non-trivial Assignment)* $P(Z = 1) \in (0, 1)$.

These assumptions are the standard IV/LATE assumptions (Imbens and Angrist, 1994; Angrist et al., 1996) with the added requirement that instrument Z is excluded from directly causing selection (2.1.1) and is allocated independently of potential selection (2.1.2). This matches classic encouragement or eligibility designs with selected samples where it is credible that the instrument causes selection and outcome only indirectly through the treatment.

Within this framework, types and causal effects can be decomposed into extensive and intensive margins. Point identification of the extensive margin effect is possible only for *compliers* $D_1 > D_0$. In particular, the extensive-margin effect for compliers is

$$\theta_E := E[S_1 - S_0|D_1 > D_0], \tag{2.1}$$

which under Assumption 2.1 is point identified by the usual Wald ratio

$$\theta_E = \frac{E[S|Z=1] - E[S|Z=0]}{E[D|Z=1] - E[D|Z=0]}. \tag{2.2}$$

Our main target parameter is the average causal effect for compliers at the intensive margin, i.e., the compliers $D_1 > D_0$ who are selected regardless of treatment response $S_1 = S_0 = 1$. We refer to this parameter as the *always-selected* or *survivor local average treatment effect* (SLATE):

$$\theta_{SLATE} := E[Y_1 - Y_0|S_1 = S_0 = 1, D_1 > D_0]. \tag{2.3}$$

Remark 1. *Without noncompliance, θ_{SLATE} reduces to the intensive margin/always-selected ATE in Lee (2009). Without sample selection, θ_{SLATE} reduces to the LATE (Imbens and Angrist, 1994). Without both, it collapses to the ATE. θ_{SLATE} is the natural intensive margin analogue of the LATE: It focuses on always-selected compliers, whose treatment status is changed by the instrument while their selection status is not. It is therefore a policy effect on the level of the outcome for a principal stratum, uncontaminated by selection into the observed sample or noncompliance. The policy relevance of θ_{SLATE} has to be discussed on a case-by-case basis. We note that, in what follows, our assumptions will be able to identify the share of always-selected compliers as well as their observable characteristics. This can be used to compare their features with those of the overall population or other relevant groups, thereby assessing how substantive the parameter is in a given empirical setting. We provide all details in Section 5 and empirical examples in Section 6 and Appendix E.*

Under selected samples θ_{SLATE} is only partially identified because the joint selection status (S_1, S_0) is not observed. In particular, even for compliers who are selected only under treatment $(S_1 = 1, S_0 = 0)$ or only under control $(S_1 = 0, S_0 = 1)$, either Y_0 or Y_1 is never observed, so their treatment effects are not identified without additional (untestable) assumptions. To make progress, we first state some standard IV identities under Assumption 2.1:

Lemma 2.1 *Suppose Assumption 2.1 holds. Then, for any integrable function $g : \mathcal{Y} \to \mathbb{R}$,*

$$P(D_1 > D_0) = E[D|Z = 1] - E[D|Z = 0], \tag{2.4}$$

$$P(S_1 = 1, D_1 > D_0) = E[DS|Z = 1] - E[DS|Z = 0], \tag{2.5}$$

$$E[g(Y_1)|S_1 = 1, D_1 > D_0] = \frac{E[DSg(Y)|Z = 1] - E[DSg(Y)|Z = 0]}{E[DS|Z = 1] - E[DS|Z = 0]}. \tag{2.6}$$

Moreover, replacing D with $(D - 1)$ in (2.5) and (2.6) identifies $P(S_0 = 1, D_1 > D_0)$ and $E[g(Y_0)|S_0 = 1, D_1 > D_0]$, respectively.

In particular, taking $g(Y) = \mathbb{1}(Y \leq y)$ yields conditional CDFs $F_{Y_1|S_1=1,D_1>D_0}(y)$ and $F_{Y_0|S_0=1,D_1>D_0}(y)$, while taking $g(Y) = Y$ yields the corresponding conditional means.

Next, we introduce the complier strata that underlie the SLATE parameter. Among compliers ($D_1 > D_0$), define

$$ac := \{S_0 = S_1 = 1,\ D_1 > D_0\} \quad \text{always-selected compliers} \tag{2.7}$$

$$dc := \{S_1 < S_0,\ D_1 > D_0\} \quad \text{treatment-de-selected compliers} \tag{2.8}$$

$$cc := \{S_1 > S_0,\ D_1 > D_0\} \quad \text{treatment-selected compliers} \tag{2.9}$$

Rewriting (2.3) in terms of these strata yields

$$\theta_{SLATE} = E[Y_1|ac] - E[Y_0|ac]. \tag{2.10}$$

The set $\{S_0 = 1, D_1 > D_0\}$ combines ac and dc, while $\{S_1 = 1, D_1 > D_0\}$ combines ac and cc. Lemma 2.1 allows us to identify $P(S_0 = 1, D_1 > D_0)$ and $P(S_1 = 1, D_1 > D_0)$, which yield two linear restrictions on the shares of three unknown types (ac, dc, cc). To fully identify these shares, we further impose a strong sample selection monotonicity condition for the compliers:

Assumption 2.2 (Strong Sample Selection Monotonicity for Compliers) *Either $P(S_1 \geq S_0|D_1 > D_0) = 1$ or $P(S_1 \leq S_0|D_1 > D_0) = 1$.*

Assumption 2.2 imposes a common direction of treatment-induced sample selection within the complier stratum $D_1 > D_0$. In particular, the treatment is assumed to either weakly increase or weakly decrease selection for all compliers. It cannot move some compliers into the sample and others out of it, effectively ruling out either *dc* or *cc* types.[4] Note that no analogous restriction is needed for strata with $D_1 = D_0$ (always-takers and never-takers) as their treatment status does not vary with the instrument. In particular, never-takers satisfy $D_0 = D_1 = 0$, so their observed selection status is governed only by S_0; always-takers satisfy $D_0 = D_1 = 1$, so their observed selection status is governed only by S_1. The counterfactual selection status under the unrealized treatment therefore does not enter the identification argument.

We now provide identification for the case where $P(S_1 \geq S_0|D_1 > D_0) = 1$. The case where $P(S_1 \leq S_0|D_1 > D_0) = 1$ can be handled analogously. Under Assumption 2.2, the stratum $S_0 = 1, D_1 > D_0$ consists only of always-selected compliers (*ac*), while $S_1 = 1, D_1 > D_0$ combines *ac* and *cc*. Lemma 2.1 then implies that

$$\beta_0 := E[Y_0|ac] = E[Y_0|S_0 = 1, D_1 > D_0] \tag{2.11}$$

is point identified by taking $g(Y) = Y$ and replacing D with $(D-1)$ in (2.6). In contrast, $E[Y_1|ac]$ is not point identified. However, Lemma 2.1 and Assumption 2.2 allow us to identify the mixture distribution of Y_1 for *ac* and *cc*,

$$F_{Y_1|ac\cup cc}(y) := F_{Y_1|S_1=1,D_1>D_0}(y), \tag{2.12}$$

as well as the frequencies of *ac* and *cc* within this mixture. Let

$$\pi_{ac} := P(S_0 = S_1 = 1, D_1 > D_0), \quad \pi_{cc} := P(S_1 > S_0, D_1 > D_0). \tag{2.13}$$

Then, under Assumption 2.2,

$$\pi_{ac} = P(S_0 = 1, D_1 > D_0), \qquad \pi_{ac} + \pi_{cc} = P(S_1 = 1, D_1 > D_0), \tag{2.14}$$

and both probabilities are identified. Moreover, define

$$p := \frac{\pi_{ac}}{\pi_{ac} + \pi_{cc}}. \tag{2.15}$$

[4] This restriction is implied, for example, by an additively separable selection equation, see, e.g., Vytlacil (2002) for treatment selection or Heiler et al. (2024) for sample selection. However, these models impose monotonicity for the full population and are thus stronger than Assumption 2.2.

p is equal to the fraction of always-selected compliers in mixture $S_1 = 1, D_1 > D_0$. Since we only observe this mixture, sharp bounds are obtained by assigning ac to the "best" or "worst" location within the mixture distribution of Y_1. Let $F_1(y) := F_{Y_1|S_1=1,D_1>D_0}(y)$ and $Q_1(u) := \inf\{y \in \mathcal{Y} : F_1(y) \geq u\}$ denote the identified CDF and quantile function of this mixture.

To simplify exposition for the remainder of this section, we maintain the following regularity condition: the relevant identified distributions entering the trimming formulas have continuous, strictly increasing CDFs. This rules out mass points at the relevant trimming cutoffs and allows the bounds to be written as ordinary trimmed means. This condition is not needed for identification per se.[5]

The lower and upper bounds for $E[Y_1|ac]$ are

$$\beta_{L,1} := E[Y_1|Y_1 \leq Q_1(p), ac \cup cc] = E[Y_1|Y_1 \leq Q_1(p), S_1 = 1, D_1 > D_0], \tag{2.16}$$

$$\beta_{U,1} := E[Y_1|Y_1 \geq Q_1(1-p), ac \cup cc] = E[Y_1|Y_1 \geq Q_1(1-p), S_1 = 1, D_1 > D_0], \tag{2.17}$$

i.e., the trimmed means obtained by placing all ac individuals in the bottom p fraction (for the lower bound) or the top p fraction (for the upper bound) of the mixture. We obtain the following proposition:

Proposition 2.1 (Sharpness) *Suppose Assumptions 2.1– 2.2 hold and $\pi_{ac} > 0$. Then,*

$$\beta_{L,1} \leq E[Y_1|ac] \leq \beta_{U,1}.$$

Moreover, $\beta_{L,1}$ and $\beta_{U,1}$ are sharp: They are, respectively, the largest lower bound and the smallest upper bound for $E[Y_1|ac]$ that are consistent with the observed data and Assumptions 2.1–2.2. Any other valid bounds under these assumptions contain $[\beta_{L,1}, \beta_{U,1}]$.

Combining the point-identified β_0 with these bounds yields sharp bounds for θ_{SLATE}:

$$\beta_L := \beta_{L,1} - \beta_0, \qquad \beta_U := \beta_{U,1} - \beta_0. \tag{2.18}$$

Using Assumption 2.2 and Lemma 2.1 with $g(Y) = Y\mathbb{1}(Y \leq Q_1(p))$ and $g(Y) = Y\mathbb{1}(Y \geq$

[5] A weaker local regularity condition would suffice: for each result below, it is enough that the corresponding identified CDF be continuous and strictly increasing in a neighborhood of the relevant trimming threshold(s). Without this regularity, the same conclusions also hold once the bounds are written in exact fractional-trimming (quantile-integral) form.

$Q_1(1-p)$), all components of β_L and β_U, namely π_{ac}, p, F_1, Q_1, and the trimmed means are functionals of the observed distribution of (SY, S, D, Z). In particular,

$$\beta_L = \frac{1}{\pi_{ac}} \left\{ \begin{array}{c} E\left[DSY\mathbb{1}(Y \le Q_1(p)) + (1-D)SY|Z=1\right] \\ -E\left[DSY\mathbb{1}(Y \le Q_1(p)) + (1-D)SY|Z=0\right] \end{array} \right\}, \tag{2.19}$$

$$\beta_U = \frac{1}{\pi_{ac}} \left\{ \begin{array}{c} E\left[DSY\mathbb{1}(Y \ge Q_1(1-p)) + (1-D)SY|Z=1\right] \\ -E\left[DSY\mathbb{1}(Y \ge Q_1(1-p)) + (1-D)SY|Z=0\right] \end{array} \right\}, \tag{2.20}$$

where

$$p = \frac{E\left[(D-1)\,S|Z=1\right] - E\left[(D-1)\,S|Z=0\right]}{E\left[DS|Z=1\right] - E\left[DS|Z=0\right]}, \tag{2.21}$$

$$Q_1(u) = \inf\{y \in \mathcal{Y} : F_1(y) \ge u\} \text{ for any } u \in (0,1)\,, \tag{2.22}$$

$$F_1(y) = \frac{E\left[\mathbb{1}\{Y \le y\}DS|Z=1\right] - E\left[\mathbb{1}\{Y \le y\}DS|Z=0\right]}{E\left[DS|Z=1\right] - E\left[DS|Z=0\right]}. \tag{2.23}$$

We next compare our bounds with two benchmarks: the ITT bounds of Lee (2009) and the CF bounds, which target θ_{SLATE} under essentially identical identification assumptions.

2.2. Comparison with Existing Bounds

2.2.1. Comparison with Lee (2009) Bounds

Without sample selection, the ITT effect

$$\tau_Y := E[Y|Z=1] - E[Y|Z=0] \tag{2.24}$$

and the average treatment effect for compliers,

$$\theta_{LATE} := \frac{\tau_Y}{\tau_D}, \qquad \tau_D := E[D|Z=1] - E[D|Z=0], \tag{2.25}$$

differ only by the scaling factor τ_D, which corresponds to the share of compliers. That is, in the standard IV point identification setting, ITT can be converted into the complier average treatment effect by dividing by τ_D. This does not generalize to treatment effect bounds under sample selection.

In particular, in the presence of sample selection, Lee (2009) derives sharp bounds for the ITT effect among the always-selected assuming strong sample selection monotonicity.

These "Lee bounds" can be viewed as a special case of our basic bounds under perfect compliance, i.e., when treatment equals instrument, $D = Z$. We only consider the lower bound. The analysis for the upper bound is analogous and omitted for brevity. Setting $D = Z$ in Equation (2.19), we obtain the bound

$$\begin{aligned}\beta_L^{Lee} &:= E[Y|Y \le Q_1^{Lee}(p^{Lee}), S = 1, Z = 1] - E[Y|S = 1, Z = 0] \\ &= \beta_{L,1}^{Lee} - \beta_0^{Lee}, \end{aligned} \tag{2.26}$$

where $Q_1^{Lee}(u) := \inf\{y \in \mathcal{Y} : F_{Y|S=1,Z=1}(y) \ge u\}$ is the u-quantile of $(Y|S = 1, Z = 1)$, and $p^{Lee} := E[S|Z = 0]/E[S|Z = 1]$ is the trimming fraction.

To relate (2.26) to our SLATE bounds, it is useful to decompose the selected population into latent types. In addition to the complier types introduced in Section 2.1, let

$$an := \{S_0 = 1, D_0 = D_1 = 0\} \quad \text{always-selected never-takers,} \tag{2.27}$$

$$aa := \{S_1 = 1, D_0 = D_1 = 1\} \quad \text{always-selected always-takers,} \tag{2.28}$$

and denote their shares by π_{an} and π_{aa} respectively. Under Assumptions 2.1–2.2, the selected group in each arm of the experiment is a mixture of these types. We show in Appendix A.2.1 that

$$F_{Y|S=1,Z=1}(y) = \omega F_{Y_0|an}(y) + (1 - \omega) F_{Y_1|ac\cup cc\cup aa}(y), \tag{2.29}$$

$$F_{Y|S=1,Z=0}(y) = \delta F_{Y_0|ac\cup an}(y) + (1 - \delta) F_{Y_1|aa}(y), \tag{2.30}$$

for suitable mixture weights $\omega, \delta \in (0, 1)$ depending only on type shares. Let $q := Q_1^{Lee}(p^{Lee})$ denote the Lee trimming threshold. The trimmed mean in the first term of (2.26) can be written as

$$\begin{aligned}\beta_{L,1}^{Lee} &:= E[Y \mid Y \le q,\ S = 1,\ Z = 1] \\ &= w(q)\, E[Y_0 \mid Y_0 \le q,\ an] + \Big(1 - w(q)\Big)\, E[Y_1 \mid Y_1 \le q,\ ac \cup cc \cup aa]\,, \end{aligned} \tag{2.31}$$

where $w(q) = \frac{\omega F_{Y_0|an}(q)}{p^{Lee}}$ is the effective an stratum weight from trimming applied after mixing. The second term of (2.26) can be written as

$$\beta_0^{Lee} := E\left[Y|S=1, Z=0\right]$$
$$= E\left[Y_0|ac \cup an\right]\delta + E\left[Y_1|aa\right](1-\delta)\,. \tag{2.32}$$

Thus, β_L^{Lee} is a difference of two mixtures that involve not only treatment compliers (ac, cc), but also always-takers and never-takers (aa, an). In contrast, our lower bound for SLATE,

$$\beta_L = E[Y_1|Y_1 \leq Q_1(p), ac \cup cc] - E[Y_0|ac], \tag{2.33}$$

depends only on the treatment complier types (ac and cc), with p equal to the fraction of always-selected compliers within that complier mixture.

Except in knife-edge cases where the shares of always-selected always-takers and never-takers vanish ($\pi_{aa} = \pi_{an} = 0$), the Y_1 term in β_L^{Lee} does not reduce to $E[Y_1|Y_1 \leq Q_1(p), ac \cup cc]$, and the Y_0 term does not reduce to $E[Y_0|ac]$. Consequently, Lee (2009) bounds cannot be transformed into sharp bounds for θ_{SLATE} by simple rescaling via primitive probabilities. We summarize this result formally in the following proposition.

Proposition 2.2 (Lee bounds do not rescale to SLATE bounds) *Suppose Assumptions 2.1–2.2 hold and $\pi_{ac} > 0$. If $\pi_{an} + \pi_{aa} > 0$, then, for generic joint distributions of (Y_0, Y_1) across latent types,*

$$\beta_L^{Lee} \neq \beta_L.$$

Moreover, there does not exist a scalar function c, depending only on the primitive probabilities (i.e. the principal-strata shares and instrument probabilities), such that

$$\beta_L = c\,\beta_L^{Lee}$$

uniformly over DGPs.

Proposition 2.2 implies that Lee (2009) ITT bounds cannot be converted into sharp bounds for SLATE by a simple rescaling analogous to τ_Y/τ_D in the standard IV case without sample selection.

2.2.2. Comparison with Chen and Flores (2015) Bounds

CF study partial identification of θ_{SLATE} under essentially identical model assumptions. We show that our bounds are weakly tighter uniformly over DGPs, and strictly tighter on a nonempty subset of DGPs than the CF bounds. For brevity, we focus on the lower bounds. The upper bounds can be treated analogously.

CF propose two lower bounds for $E[Y_1|ac]$, based on trimming the distribution of Y among selected treated units, $Y|DSZ = 1$, equivalent to $Y_1|ac \cup cc \cup aa$ under Assumptions 2.1–2.2, and then taking the maximum of the two. In contrast, our sharp bounds trim within the smaller identified mixture $Y_1|ac \cup cc$. Let $Q_1^{CF}(u)$ denote the u-quantile of $Y|DSZ = 1$. Their basic lower bound is

$$\begin{aligned} \beta_L^{mix} &:= E\left[Y|DSZ = 1, Y \leq Q_1^{CF}(p_1^{CF})\right] - \beta_0 \\ &= \beta_{L,1}^{mix} - \beta_0, \end{aligned} \tag{2.34}$$

where $p_1^{CF} := \pi_{ac}/(\pi_{ac} + \pi_{cc} + \pi_{aa})$ and $\beta_0 := E[Y_0|ac]$. Their alternative lower bound uses additional information from the always-taker stratum (aa),

$$\begin{aligned} \beta_L^{adj} &:= \left\{ \begin{array}{l} \left(1 + \dfrac{\pi_{aa}}{\pi_{ac}}\right) E\left[Y \mid DSZ = 1, Y \leq Q_1^{CF}(p_2^{CF})\right] \\ -\dfrac{\pi_{aa}}{\pi_{ac}} E\left[Y \mid DS(1-Z) = 1\right] \end{array} \right\} - \beta_0 \\ &= \beta_{L,1}^{adj} - \beta_0, \end{aligned} \tag{2.35}$$

where $p_2^{CF} := (\pi_{ac} + \pi_{aa})/(\pi_{ac} + \pi_{cc} + \pi_{aa})$ and $E[Y|DS(1-Z) = 1] = E[Y_1|aa]$ is point identified. The CF lower bound is then

$$\beta_L^{CF} := \max\{\beta_L^{mix}, \beta_L^{adj}\}. \tag{2.36}$$

Under Assumptions 2.1–2.2, the conditional distribution $Y|DSZ = 1$ coincides with Y_1 for the mixture of types ac, cc, and aa:

$$F_{Y|DSZ=1}(y) = F_{Y_1|ac \cup cc \cup aa}(y), \tag{2.37}$$

so both β_L^{mix} and β_L^{adj} are obtained by trimming this mixture distribution and then reweighting to isolate ac and aa. CF show that β_L^{mix} corresponds to the worst-case lower bound when all ac individuals are placed in the bottom p_1^{CF}-fraction of $Y_1|ac \cup cc \cup aa$, while β_L^{adj} corresponds to the worst case when ac and aa together occupy the bottom p_2^{CF}-fraction of that distribution, using the fact that $E[Y_1|aa]$ is identified.

In contrast, our lower bound β_L uses the smallest stratum that is point identified, namely $Y_1|ac \cup cc$. As shown in Lemma A.2 in Appendix A.3, the distribution $F_{Y_1|ac\cup cc}$ can be recovered from the two observable selected distributions

$$F_{Y|DSZ=1}(y) = F_{Y_1|ac\cup cc\cup aa}(y), \qquad F_{Y|DS(1-Z)=1}(y) = F_{Y_1|aa}(y). \tag{2.38}$$

We then construct the sharp lower bound

$$\beta_L = E\left[Y_1|Y_1 \leq Q_1(p), ac \cup cc\right] - \beta_0, \qquad p := \frac{\pi_{ac}}{\pi_{ac} + \pi_{cc}}, \tag{2.39}$$

by trimming Y_1 only within this complier mixture.

Intuitively, β_L assumes, in a worst-case fashion, that all ac individuals are located below all cc individuals in the distribution of Y_1 among compliers, but it imposes no restrictions on the relative ranking of ac versus aa. By contrast, the CF bounds are based on worst-case restrictions on the joint ranking of ac, cc, and aa within $Y_1|ac \cup cc \cup aa$. This makes them conservative. In particular, we obtain the following proposition:

Proposition 2.3 (Dominance over CF bounds) *Suppose Assumptions 2.1–2.2 hold and $\pi_{ac} > 0$. Then,*

$$\beta_{L,1} \geq \beta_{L,1}^{CF} \geq \beta_{L,1}^{mix} \quad \textit{and} \quad \beta_{U,1} \leq \beta_{U,1}^{CF} \leq \beta_{U,1}^{mix},$$

where $\beta_{L,1}$ is defined in (2.16), $\beta_{L,1}^{mix}$ in (2.34), and $\beta_{L,1}^{CF} := \max\{\beta_{L,1}^{mix}, \beta_{L,1}^{adj}\}$, with $\beta_{L,1}^{adj}$ defined in (2.35). The upper-bound counterparts $\beta_{U,1}$, $\beta_{U,1}^{mix}$, $\beta_{U,1}^{adj}$, and $\beta_{U,1}^{CF}$ are defined analogously. Hence, imposing either of the two restrictions used by CF cannot tighten our sharp lower or upper bound for $E[Y_1 \mid ac]$.
Moreover, these inequalities are strict on a nonempty class of DGPs:

$$\beta_{L,1} > \beta_{L,1}^{CF} \iff \pi_{aa} > 0 \quad \textit{and} \quad 0 < F_{Y_1|aa}\Big(Q_1(p)\Big) < 1,$$
$$\beta_{U,1} < \beta_{U,1}^{CF} \iff \pi_{aa} > 0 \quad \textit{and} \quad 0 < F_{Y_1|aa}\Big(Q_1(1-p)\Big) < 1.$$

3. Bounds with Covariates

We now extend the unconditional baseline bounds in Section 2.1 to incorporate predetermined covariates. Incorporating such covariates serves three purposes: First, it allows for an unconfounded instead of a fully randomly assigned instrument. Second, even under fully randomized assignment, it can improve efficiency by conditioning on relevant covariates when constructing trimmed means. Third, it allows us to relax strong sample selection monotonicity. In particular, we permit the direction of sample selection for treatment compliers to vary with covariates. For simplicity, we maintain the strong treatment response monotonicity assumption 2.1.3. This no-defiers restriction is natural in eligibility and encouragement designs such as JC and OHIE: departures from perfect compliance are largely one-sided, arising mainly from incomplete take-up among those assigned to treatment rather than from substantial treatment receipt among controls. Relaxing this assumption is possible via further covariate partitioning analogous to the sample selection case in what follows.

Let $X \in \mathcal{X} \subset \mathbb{R}^{d_x}$ denote a vector of predetermined covariates. The observed data are (SY, S, D, Z, X). We impose the following conditional IV assumptions:

Assumption 3.1 (Conditional IV) *For $d = 0, 1$:*

3.1.1 (Exclusion) $Y_{d,1} = Y_{d,0} =: Y_d$, $S_{d,1} = S_{d,0} =: S_d$.

3.1.2 (Independence) $(Y_1, Y_0, S_1, S_0, D_1, D_0) \perp Z | X$.

3.1.3 (Strong Treatment Response Monotonicity) $P(D_1 \geq D_0) = 1$.

3.1.4 (First Stage) $P(D_1 \neq D_0 \mid X = x) > 0$ *for P_X-almost every $x \in \mathcal{X}$*

3.1.5 (Non-trivial Assignment) $P(Z = 1 \mid X = x) \in (0, 1)$ *for P_X-almost every $x \in \mathcal{X}$.*

These assumptions are standard extensions of the conditional IV/LATE assumptions (Frölich, 2007) generalizing Assumption 2.1 with instrument Z again being excluded from directly causing selection and being independently allocated of potential selection given covariates.

It is important to note that Assumptions 3.1.4 and 3.1.5 are not intended to impose substantive restrictions beyond their unconditional counterparts Assumptions 2.1.4 and 2.1.5. Rather, they should be interpreted as support restrictions for the conditional analysis: if the conditional first-stage or non-trivial assignment condition fails on a set of positive P_X-probability, $\mathcal{X}$ should be understood as restricted to the support on which these conditions hold.[6]

Under Assumption 3.1, the identities in Lemma 2.1 hold pointwise in x. Next, we relax strong sample selection monotonicity to a weaker conditional version that allows the direction of monotonicity to vary with covariates.

Assumption 3.2 (Weak Sample Selection Monotonicity for Compliers) *Either* $P(S_1 \geq S_0|D_1 > D_0, X = x) = 1$ *or* $P(S_1 \leq S_0|D_1 > D_0, X = x) = 1$.

Assumption 3.2 yields subsets of the covariate space where treatment compliers can have either a weakly positive or negative selection responses to treatment. At their intersection, treatment does not affect selection. Without loss of generality, we now define a distinct partitioning where this intersection is combined with the weakly positive responders. Formally, let

$$\mathcal{X}^0 = \{x \in \mathcal{X} : P(S_1 = S_0|D_1 > D_0, X = x) = 1\}$$

and partitions

$$\begin{aligned} \mathcal{X}^+ = \{x \in \mathcal{X} : P(S_1 \geq S_0|D_1 > D_0, X = x) = 1, \\ P(S_1 > S_0|D_1 > D_0, X = x) > 0\} \cup \mathcal{X}^0, \end{aligned} \tag{3.1}$$

$$\begin{aligned} \mathcal{X}^- = \{x \in \mathcal{X} : P(S_1 \leq S_0|D_1 > D_0, X = x) = 1, \\ P(S_1 < S_0|D_1 > D_0, X = x) > 0\}. \end{aligned} \tag{3.2}$$

We note that the presence of $\mathcal{X}^0$ is harmless for identification. For regular semiparametric inference, however, it will be necessary that $P(\mathcal{X}^0) = 0$ (Heiler et al., 2024). We return to this in Section 4.

Remark 2. *Assumption 3.2 weakens strong sample selection monotonicity by allowing the sign of the treatment effect on selection to vary with observed covariates. Thus, treatment may increase selection for some values in $\mathcal{X}$ and decrease it for others. The restriction is nevertheless substantive: conditional on $X = x$ and complier status, residual heterogeneity in the selection response must be one-sided. In particular, the assumption rules out the coexistence of treatment-induced entry and exit among compliers with the*

[6] Failure of the first stage and failure of non-trivial assignment have slightly different interpretations. If $P(D_1 \neq D_0 \mid X = x) = 0$, there are no compliers at that x. so complier-based conditional objects are not meaningful there. If $P(Z = 1 \mid X = x) \in \{0, 1\}$, there is no within-$x$ variation in the instrument, so the conditional IV comparisons are not identified there.

same covariate values. Its credibility therefore depends on whether $\mathcal{X}$ is rich enough to absorb the economically relevant heterogeneity in the direction of selection. It is more plausible when the main determinants of the sign of the selection response are observed, such as baseline characteristics governing participation, employment, survival, etc. It is less plausible when unobserved factors can generate opposing responses within the same covariate cell, see Heiler et al. (2024) and Semenova (2025) for additional discussion.

Let $\pi_{ac}(x) := P(ac|X = x)$ denote the conditional share of always-selected compliers, and let $\pi_{ac} := E[\pi_{ac}(X)]$ denote their unconditional share. Further let $\mathcal{X}_{ac} := \{x \in \mathcal{X} : \pi_{ac}(x) > 0\}$.[7] For any $x \in \mathcal{X}_{ac}$, define the conditional SLATE parameter

$$\theta_{SLATE}(x) := E[Y_1 - Y_0|ac, X = x]. \tag{3.3}$$

Let

$$\lambda_d(x) := P(S_d = 1, D_1 > D_0|X = x), \quad d = 0, 1, \tag{3.4}$$

and write $\pi_{cc}(x) := P(cc|X = x)$ and $\pi_{dc}(x) := P(dc|X = x)$ for the conditional shares of cc and dc, respectively. Under Assumption 3.2,

$$\lambda_0(x) = \pi_{ac}(x), \quad \lambda_1(x) = \pi_{ac}(x) + \pi_{cc}(x) \quad \text{for } x \in \mathcal{X}^+, \tag{3.5}$$

$$\lambda_1(x) = \pi_{ac}(x), \quad \lambda_0(x) = \pi_{ac}(x) + \pi_{dc}(x) \quad \text{for } x \in \mathcal{X}^-. \tag{3.6}$$

This yields the ratio

$$p(x) := \frac{\lambda_0(x)}{\lambda_1(x)}. \tag{3.7}$$

Thus $0 \leq p(x) \leq 1$ on $\mathcal{X}_{ac}^+ := \mathcal{X}^+ \cap \mathcal{X}_{ac}$ and $p(x) > 1$ on $\mathcal{X}_{ac}^- := \mathcal{X}^- \cap \mathcal{X}_{ac}$. For $d = 0, 1$ and $x \in \mathcal{X}_{ac}$, define the conditional quantile

$$Q_d(u, x) := \inf \left\{ y \in \mathcal{Y} : F_{Y_d|S_d=1, D_1>D_0, X=x}(y) \geq u \right\}. \tag{3.8}$$

To simplify exposition, for the remainder of this section we again maintain that the relevant identified conditional CDF entering each trimming formula is continuous and strictly increasing in a neighborhood of the corresponding trimming threshold(s). This

[7] For identification of the unconditional bounds, the substantive requirement is $\pi_{ac} > 0$. $\pi_{ac}(x)$ may be zero on subsets of the covariate support as they receive zero weight in numerator and denominator of the unconditional bounds. As a convention for display, we treat $\theta_{SLATE}(x) = 0$ for $x \in \mathcal{X} \setminus \mathcal{X}_{ac}$.

rules out mass points at the trimming thresholds and allows the conditional sharp bounds below to be written as ordinary trimmed means.[8] Then Proposition 2.1 can be applied pointwise in x. This yields the following sharp lower and upper bounds for $\theta_{SLATE}(x)$:

$$\begin{aligned}\beta_L^+(x) &:= E\left[Y_1|S_1=1, D_1>D_0, Y_1 \leq Q_1(p(x),x), X=x\right] \\ &\quad - E\left[Y_0|S_0=1, D_1>D_0, X=x\right], \qquad (3.9)\\ \beta_U^+(x) &:= E\left[Y_1|S_1=1, D_1>D_0, Y_1 \geq Q_1(1-p(x),x), X=x\right] \\ &\quad - E\left[Y_0|S_0=1, D_1>D_0, X=x\right], \qquad (3.10)\end{aligned}$$

for $x \in \mathcal{X}_{ac}^+$, and

$$\begin{aligned}\beta_L^-(x) &:= E\left[Y_1|S_1=1, D_1>D_0, X=x\right] \\ &\quad - E\left[Y_0|S_0=1, D_1>D_0, Y_0 \geq Q_0(1-1/p(x),x), X=x\right], \qquad (3.11)\\ \beta_U^-(x) &:= E\left[Y_1|S_1=1, D_1>D_0, X=x\right] \\ &\quad - E\left[Y_0|S_0=1, D_1>D_0, Y_0 \leq Q_0(1/p(x),x), X=x\right], \qquad (3.12)\end{aligned}$$

for $x \in \mathcal{X}_{ac}^-$. As a notational convention, set $\beta_B^\pm(x) = 0$ for $x \in \mathcal{X} \setminus \mathcal{X}_{ac}$, $B \in \{L, U\}$. These values are used only in the reduced-forms $\beta_B^\pm(x)\pi_{ac}(x)$ entering the unconditional bounds. We note that when treatment does not cause selection, conditional effects bounds collapse to a point, i.e. for any $x \in (\mathcal{X}^0 \cap \mathcal{X}_{ac}) \subseteq \mathcal{X}_{ac}^+$ we have that

$$\beta_L^+(x) = \beta_U^+(x) = E\left[Y_1|S_1=1, D_1>D_0, X=x\right] - E\left[Y_0|S_0=1, D_1>D_0, X=x\right]. \qquad (3.13)$$

Now let

$$\mathbb{1}^+(x) := \mathbb{1}(x \in \mathcal{X}^+), \qquad \mathbb{1}^-(x) := \mathbb{1}(x \in \mathcal{X}^-), \qquad (3.14)$$

and define for any $B \in \{L, U\}$

$$\beta_B(x) := \mathbb{1}^+(x)\beta_B^+(x) + \mathbb{1}^-(x)\beta_B^-(x). \qquad (3.15)$$

[8] For $x \in \mathcal{X}_{ac}^+$, the relevant CDF is $F_{Y_1|S_1=1,D_1>D_0,X=x}(\cdot)$ and the relevant thresholds are $Q_1(p(x),x)$ and $Q_1(1-p(x),x)$. For $x \in \mathcal{X}_{ac}^-$, the relevant CDF is $F_{Y_0|S_0=1,D_1>D_0,X=x}(\cdot)$ and the relevant thresholds are $Q_0(1-1/p(x),x)$ and $Q_0(1/p(x),x)$. Without this regularity, the same bounds can be written in exact fractional-trimming form.

The unconditional SLATE parameter can be written as

$$\theta_{SLATE} = \frac{E[\theta_{SLATE}(X)\pi_{ac}(X)]}{\pi_{ac}}, \tag{3.16}$$

The corresponding sharp lower and upper bounds are

$$\beta_L = \frac{E[\beta_L(X)\pi_{ac}(X)]}{\pi_{ac}}, \qquad \beta_U = \frac{E[\beta_U(X)\pi_{ac}(X)]}{\pi_{ac}}. \tag{3.17}$$

To make the link with observed data explicit, we next provide the estimands of the denominator and numerator moment functions in β_L and β_U. Define the instrument propensity score $e(x) := P(Z = 1|X = x)$, and the inverse probability weight

$$W(Z, X) := \frac{Z}{e(X)} - \frac{1 - Z}{1 - e(X)}. \tag{3.18}$$

Extending Lemma 2.1 to hold conditional on X, one can show that

$$\begin{aligned} \pi_{ac} &= E\left[\mathbb{1}^+(X)\lambda_0(X) + \mathbb{1}^-(X)\lambda_1(X)\right] \\ &= E\left[\min\{\lambda_0(X), \lambda_1(X)\}\right]. \end{aligned} \tag{3.19}$$

where now

$$\lambda_0(x) = E\left[W(Z, X)(D - 1)S|X = x\right], \tag{3.20}$$

$$\lambda_1(x) = E\left[W(Z, X)DS|X = x\right], \tag{3.21}$$

$$\mathbb{1}^+(x) = \mathbb{1}(\lambda_0(x) \leq \lambda_1(x)), \tag{3.22}$$

$$\mathbb{1}^-(x) = \mathbb{1}(\lambda_0(x) > \lambda_1(x)). \tag{3.23}$$

Thus π_{ac} is identified. In addition, for any y and $x \in \mathcal{X}_{ac}$,

$$F_1(y|x) := F_{Y_1|S_1=1,D_1>D_0,X=x}(y) = \frac{E\left[W(Z, X)DS\mathbb{1}(Y \leq y)|X = x\right]}{E\left[W(Z, X)DS|X = x\right]}, \tag{3.24}$$

$$F_0(y|x) := F_{Y_0|S_0=1,D_1>D_0,X=x}(y) = \frac{E\left[W(Z, X)(D - 1)S\mathbb{1}(Y \leq y)|X = x\right]}{E\left[W(Z, X)(D - 1)S|X = x\right]}. \tag{3.25}$$

The conditional quantiles $Q_d(u, x)$ used for trimming in the conditional bounds in Equations (3.9) – (3.12) are defined as the generalized inverses of these CDFs and are thus identified. Using these identities, the following proposition collects the explicit estimands for the unconditional bounds β_L and β_U.

Proposition 3.1 *Suppose Assumptions 3.1–3.2 hold and $\pi_{ac} > 0$. Then the sharp lower and upper bounds for θ_{SLATE} are identified as*

$$\beta_L = \frac{E[\beta_L(X)\pi_{ac}(X)]}{\pi_{ac}}, \qquad \beta_U = \frac{E[\beta_U(X)\pi_{ac}(X)]}{\pi_{ac}},$$

where π_{ac} is given by Equation (3.19) and

$$\beta_L(x) := \mathbb{1}^+(x)\beta_L^+(x) + \mathbb{1}^-(x)\beta_L^-(x),$$
$$\beta_U(x) := \mathbb{1}^+(x)\beta_U^+(x) + \mathbb{1}^-(x)\beta_U^-(x),$$

with

$$\beta_L^+(x)\pi_{ac}(x) := E\left[W(Z,X)\left(DSY\mathbb{1}(Y \le Q_1(p(x),x)) + (1-D)SY\right)|X = x\right], \tag{3.26}$$
$$\beta_L^-(x)\pi_{ac}(x) := E\left[W(Z,X)\left(DSY + (1-D)SY\mathbb{1}(Y \ge Q_0(1-1/p(x),x))\right)|X = x\right], \tag{3.27}$$
$$\beta_U^+(x)\pi_{ac}(x) := E\left[W(Z,X)\left(DSY\mathbb{1}(Y \ge Q_1(1-p(x),x)) + (1-D)SY\right)|X = x\right], \tag{3.28}$$
$$\beta_U^-(x)\pi_{ac}(x) := E\left[W(Z,X)\left(DSY + (1-D)SY\mathbb{1}(Y \le Q_0(1/p(x),x))\right)|X = x\right], \tag{3.29}$$

on $\mathcal{X}_{ac}$, and $\beta_B^{\pm}(x)\pi_{ac}(x) = 0$ by convention on $\mathcal{X} \setminus \mathcal{X}_{ac}$. $\lambda_d(x)$ and $W(z,x)$ are point identified on $\mathcal{X}$ and conditional quantiles $Q_d(\cdot,x)$ are point identified on $\mathcal{X}_{ac}$.

In summary, (β_L, β_U) are functionals of the observed law of (SY, S, D, Z, X) and a collection of nuisance functions (instrument propensity score, conditional means, and (inverted) conditional distribution functions). In the next section, we derive orthogonal moment conditions and semiparametric influence functions for these functionals, which will allow us to construct semiparametrically efficient debiased machine learning estimators for (β_L, β_U) and confidence intervals for θ_{SLATE}.

4. Estimation and Inference

4.1. Nuisance Functions and Target Parameter

The previous section shows that for any $B \in \{L, U\}$, the unconditional bound can be written as

$$\beta_B = \frac{E[\beta_B(X)\pi_{ac}(X)]}{\pi_{ac}}, \tag{4.1}$$

with $\beta_B(X)$ defined in Proposition 3.1 and π_{ac} in (3.19). For notational convenience, we decompose the numerator of (4.1) as

$$E[\beta_B(X)\pi_{ac}(X)] = N_0^+ + N_0^- + N_{B,1}^+ + N_{B,1}^-, \tag{4.2}$$

where, for $B \in \{L, U\}$ and $* \in \{+, -\}$, we define

$$N_0^* := E[\beta_0^*(X)\pi_{ac}(X)], \qquad N_{B,1}^* := E[\beta_{B,1}^*(X)\pi_{ac}(X)]. \tag{4.3}$$

Table 4.1 contains all primary nuisance functions as well as all derived quantities that are used for the remainder of this section. We collect the required primary nuisance functions in vector

$$\eta(x) := (e(x), \{r(z,x), m(z,x), \mu(z,x), \nu(z,x), F(\cdot|d,z,x), G_L(\cdot|d,z,x), G_U(\cdot|d,z,x)\}_{d,z}). \tag{4.4}$$

4.2. Semiparametric Influence Functions and Nuisance Functions

For any $B \in \{L, U\}$, we treat π_{ac} in (3.19) as well as the four components in (4.2) as target functionals and derive their semiparametric influence functions/efficient orthogonal moments. They are then combined to yield the efficient influence function and estimator for the respective β_B. Denote the influence function operator $\mathbb{IF} : \Theta \to L^2(P, \mathbb{R}^q)$ as the operator that, for a parameter $\theta : \mathcal{P} \to \mathbb{R}^q$, returns its influence function, i.e., the standard

Table 4.1: Primary and Derived Nuisance Functions for Bounds (β_L, β_U)

Primary Nuisance Parameter in η	Definition
$e(x)$	$P(Z=1\|X=x)$
$r(z,x)$	$E[DS\|Z=z,X=x]$
$m(z,x)$	$E[(1-D)S\|Z=z,X=x]$
$\mu(z,x)$	$E[(1-D)SY\|Z=z,X=x]$
$\nu(z,x)$	$E[DSY\|Z=z,X=x]$
$F(y\|d,z,x)$	$E[\mathbb{1}(Y \leq y)\|D=d,S=1,Z=z,X=x]$
$G_L(y\|d,z,x)$	$E[Y\mathbb{1}(Y \leq y)\|D=d,S=1,Z=z,X=x]$
$G_U(y\|d,z,x)$	$E[Y\mathbb{1}(Y > y)\|D=d,S=1,Z=z,X=x]$
Derived Nuisance Parameter or Variable	**Definition**
$\lambda_0(x)$	$-(m(1,x)-m(0,x))$
$\lambda_1(x)$	$r(1,x)-r(0,x)$
$\pi_{ac}(x)$	$\min\{\lambda_0(x),\lambda_1(x)\}$
$p(x)$	$\frac{\lambda_0(x)}{\lambda_1(x)}$
$\mathbb{1}^+(x)$	$\mathbb{1}(p(x) \leq 1)$
$\mathbb{1}^-(x)$	$\mathbb{1}(p(x) > 1)$
$W(z,x)$	$\frac{z}{e(x)} - \frac{1-z}{1-e(x)}$
$F_1(y\|x)$	$\frac{F(y\|1,1,x)r(1,x)-F(y\|1,0,x)r(0,x)}{r(1,x)-r(0,x)}$
$F_0(y\|x)$	$\frac{F(y\|0,1,x)m(1,x)-F(y\|0,0,x)m(0,x)}{m(1,x)-m(0,x)}$
$Q_1(u,x)$	$\inf\{y \in \mathcal{Y} : F_1(y\|x) \geq u\}$
$Q_0(u,x)$	$\inf\{y \in \mathcal{Y} : F_0(y\|x) \geq u\}$
$\psi_L^+(z,x)$	$G_L(Q_1(p(X),X)\|1,z,x)$
$\psi_U^+(z,x)$	$G_U(Q_1(1-p(X),X)\|1,z,x)$
$\psi_L^-(z,x)$	$G_L(Q_0(1-1/p(X),X)\|0,z,x)$
$\psi_U^-(z,x)$	$G_U(Q_0(1/p(X),X)\|0,z,x)$

Riesz-representer of the pathwise derivative (Bickel et al., 1993; Kennedy, 2024).[9]

We now present the influence function for the target parameters β_B for $B \in \{L, U\}$. We suppress dependence on nuisances and data in what follows whenever it does not cause confusion. By linearity of the $\mathbb{IF}$ operator, we have that

$$\mathbb{IF}(\beta_B) = \frac{1}{\pi_{ac}}\left[\mathbb{IF}(N_0^+) + \mathbb{IF}(N_0^-) + \mathbb{IF}(N_{B,1}^+) + \mathbb{IF}(N_{B,1}^-) - \beta_B \mathbb{IF}(\pi_{ac})\right] \tag{4.5}$$

The influence functions of the components are in Table 4.2. Putting upper and lower

[9]The influence function is defined via the functional derivative of the target parameter with respect to perturbations along the tangent space evaluated at the true probability distribution. For example, if we observe iid data $X \sim \mathcal{P}$ and the target functional is $E[X]$, then for $\mathbb{IF}(E[X]) = X - E[X]$. For a regular parameter under the nonparametric model $\mathcal{P}$, this is the unique semiparametric efficient influence function, see, e.g., Hines et al. (2022), Kennedy (2024) or Heiler and Knaus (2026) for additional examples.

Table 4.2: Influence Functions of Components

Component	Influence Function $\mathbb{IF}$
π_{ac}	$-\mathbb{1}^{+}(X)\Big[W(Z,X)((1-D)S-m(Z,X))+m(1,X)-m(0,X)\Big]$
	$+\mathbb{1}^{-}(X)\Big[W(Z,X)(DS-r(Z,X))+r(1,X)-r(0,X)\Big]-\pi_{ac}$
N_0^{+}	$\mathbb{1}^{+}(X)\Big\{W(Z,X)((1-D)SY-\mu(Z,X))+\mu(1,X)-\mu(0,X)\Big\}-N_0^{+}$
$N_{L,1}^{+}$	$\mathbb{1}^{+}(X)\Big\{-Q_1(p(X),X)W(Z,X)\Big[(1-D)S-m(Z,X)$
	$+(DS-r(Z,X))F(Q_1(p(X),X)\vert 1,Z,X)$
	$+DS(\mathbb{1}(Y\leq Q_1(p(X),X))-F(Q_1(p(X),X)\vert 1,Z,X))\Big]$
	$+W(Z,X)\Big[DSY\mathbb{1}(Y\leq Q_1(p(X),X))-\psi_L^{+}(Z,X)r(Z,X)\Big]$
	$+\psi_L^{+}(1,X)r(1,X)-\psi_L^{+}(0,X)r(0,X)\Big\}-N_{L,1}^{+}$
$N_{U,1}^{+}$	$\mathbb{1}^{+}(X)\Big\{-Q_1(1-p(X),X)W(Z,X)\Big[(1-D)S-m(Z,X)$
	$+(DS-r(Z,X))(1-F(Q_1(1-p(X),X)\vert 1,Z,X))$
	$+DS(\mathbb{1}(Y>Q_1(1-p(X),X))-(1-F(Q_1(1-p(X),X)\vert 1,Z,X)))\Big]$
	$+W(Z,X)\Big[DSY\mathbb{1}(Y>Q_1(1-p(X),X))-\psi_U^{+}(Z,X)r(Z,X)\Big]$
	$+\psi_U^{+}(1,X)r(1,X)-\psi_U^{+}(0,X)r(0,X)\Big\}-N_{U,1}^{+}$
N_0^{-}	$\mathbb{1}^{-}(X)\Big\{W(Z,X)(DSY-\nu(Z,X))+\nu(1,X)-\nu(0,X)\Big\}-N_0^{-}$
$N_{L,1}^{-}$	$\mathbb{1}^{-}(X)\Big\{-Q_0(1-\frac{1}{p(X)},X)W(Z,X)\Big[DS-r(Z,X)$
	$+((1-D)S-m(Z,X))(1-F(Q_0(1-\frac{1}{p(X)},X)\vert 0,Z,X))$
	$+(1-D)S(\mathbb{1}(Y\geq Q_0(1-\frac{1}{p(X)},X))-(1-F(Q_0(1-\frac{1}{p(X)},X)\vert 0,Z,X)))\Big]$
	$+W(Z,X)\Big[(1-D)SY\mathbb{1}(Y\geq Q_0(1-\frac{1}{p(X)},X))-\psi_L^{-}(Z,X)m(Z,X)\Big]$
	$+\psi_L^{-}(1,X)m(1,X)-\psi_L^{-}(0,X)m(0,X)\Big\}-N_{L,1}^{-}$
$N_{U,1}^{-}$	$\mathbb{1}^{-}(X)\Big\{-Q_0(\frac{1}{p(X)},X)W(Z,X)\Big[DS-r(Z,X)$
	$+((1-D)S-m(Z,X))F(Q_0(\frac{1}{p(X)},X)\vert 0,Z,X)$
	$+(1-D)S(\mathbb{1}(Y\leq Q_0(\frac{1}{p(X)},X))-F(Q_0(\frac{1}{p(X)},X)\vert 0,Z,X))\Big]$
	$+W(Z,X)\Big[(1-D)SY\mathbb{1}(Y\leq Q_0(\frac{1}{p(X)},X))-\psi_U^{-}(Z,X)m(Z,X)\Big]$
	$+\psi_U^{-}(1,X)m(1,X)-\psi_U^{-}(0,X)m(0,X)\Big\}-N_{U,1}^{-}$

bounds together then yields

$$\mathbb{IF}(\beta) = \begin{pmatrix} \mathbb{IF}(\beta_L) \\ \mathbb{IF}(\beta_U) \end{pmatrix} \tag{4.6}$$

The influence functions $\mathbb{IF}(\beta_B)$ for $B \in \{L, U\}$ nest the existing literature on LATE and sample selection bounds. In particular, we obtain the following proposition:

Proposition 4.1 *For any $B \in \{L, U\}$, the efficient influence function in* (4.5) *under (i) perfect compliance, (ii) no sample selection or (iii) both collapse to their respective efficient influence functions for Lee bounds, LATE, or ATE:*

(i) If $P(D = Z) = 1$, then $\mathbb{IF}(\beta_B) = \mathbb{IF}(\beta_B^{Lee})$ a.s.

(ii) If $P(S = 1) = 1$, then $\mathbb{IF}(\beta_B) = \mathbb{IF}(\theta_{LATE})$ a.s.

(iii) If $P(S = 1, D = Z) = 1$, then $\mathbb{IF}(\beta_B) = \mathbb{IF}(\theta_{ATE})$ a.s.

Practically, this means that, as $P(D = Z) \to 1$ or $P(S = 1) \to 1$, our influence functions will get closer in probability to the semiparametrically efficient influence functions for Lee-bounds on the intensive margin ATE (Heiler et al., 2024; Semenova, 2025) or the LATE (Frölich, 2007) respectively. If both apply, the functions approach the efficient influence of the ATE (Hahn, 1998).

4.3. Finite Sample Implementation

We now outline the steps to estimate bounds and effect confidence intervals and provide more details regarding nuisance function estimation. For a generic random variable X we denote $E_n[X] = \frac{1}{n}\sum_i^n X_i$ in what follows. Assume we have independent data $O_i = (S_iY_i, S_i, D_i, Z_i, X_i)'$ for $i = 1, \dots, n$. Algorithm 1 contains a step-by-step explanation of how to obtain bounds and inference. Estimation of bounds is fairly standard within the DML framework for composite ratio parameters. In particular, we estimate all their components separately with cross-fitted nuisances to obtain an estimate for the eventual bounds and their respective influence functions. The latter then yield the full variance-covariance matrix estimates that can be used for standard-normal-based inference on bounds or, more importantly, confidence intervals for the effect using refined critical values (Imbens and Manski, 2004; Stoye, 2020). All of these have at least $(1 - \alpha)$ asymptotic coverage under some regularity conditions on the conditional distribution and suitable rate conditions on the nuisances, see Section 4.4 for more details.

Algorithm 1 Estimation and Inference

Require: Data $\{O_i\}_{i=1}^n$, number of folds K.

Randomly partition data into K disjoint folds $I_1, \ldots, I_K$ of approximately equal size.
for each $f \in \{1, \ldots, K\}$ **do**
    estimate primary nuisance parameters using data in I_f^c.
    evaluate the nuisances on I_f.
end for
for each $N \in \{N_0^+, N_0^-, N_{L,1}^+, N_{L,1}^-, N_{U,1}^+, N_{U,1}^-, \pi_{ac}\}$ **do**
    for all $i = 1, \ldots, n$ plug in all nuisances $\hat{\eta}$ into the orthogonal moments $\mathbb{IF}(\cdot)$.
    obtain $\hat{N}$ by setting the sample mean of the orthogonal moments equal to zero

$$E_n[\mathbb{IF}(\hat{N}, \hat{\eta})] = 0$$

end for
return the bound estimators as

$$\hat{\beta}_B = \frac{\hat{N}_0^+ + \hat{N}_0^- + \hat{N}_{B,1}^+ + \hat{N}_{B,1}^-}{\hat{\pi}_{ac}}$$

return standard errors via estimated composite moment:

$$\hat{\sigma}_{B,n} = \sqrt{\frac{E_n[\mathbb{IF}(\hat{\beta}_B, \hat{\eta})^2]}{n}}$$

where the composite moment is given by

$$\mathbb{IF}(\hat{\beta}_B, \hat{\eta}) = \frac{1}{\hat{\pi}_{ac}} \left[\mathbb{IF}(\hat{N}_0^+, \hat{\eta}) + \mathbb{IF}(\hat{N}_0^-, \hat{\eta}) + \mathbb{IF}(\hat{N}_{B,1}^+, \hat{\eta}) + \mathbb{IF}(\hat{N}_{B,1}^-, \hat{\eta}) - \hat{\beta}_B \mathbb{IF}(\hat{\pi}_{ac}, \hat{\eta}) \right]$$

return $(1 - \alpha)$ effect confidence intervals as

$$CI_{1-\alpha}(\theta_{SLATE}) = \left[\hat{\beta}_L - c_{L,\alpha} \hat{\sigma}_{L,n}, \ \hat{\beta}_U + c_{U,\alpha} \hat{\sigma}_{U,n} \right]$$

where refined critical values $c_{L,\alpha}, c_{U,\alpha} \leq z_{1-\alpha/2}$ can be chosen according to Imbens and Manski (2004) or Stoye (2020). Using $z_{1-\alpha/2}$ is valid for inference on the bounds.

We now discuss primary nuisance function estimation and how to obtain derived nuisances as defined in Table 4.1. Our high-level assumptions in Section 4.4 match these primitive objects that are conditional means/probabilities, conditional CDFs, and trimmed conditional means. For primary nuisance functions that are simple conditional means or probabilities, e, r, m, μ, ν, a plethora of off-the-shelf nonparametric and machine learning methods such as neural networks, forests or high-dimensional sparse parametric models are available. The derived nuisances $\lambda_0, \lambda_1, \pi_{ac}, p, \mathbb{1}^+, \mathbb{1}^-, W$ only depend on the primary and can be obtained via simple plug-in versions.

The functional parameters F and G_B require special attention. In particular, we suggest to estimate the primary conditional CDF F via machine learning analogues of distributional regression (Foresi and Peracchi, 1995; Klein, 2024). We also recommend direct imposition or post-processing, e.g., via isotonic regression (Henzi et al., 2021) that further refine these estimates by enforcing nondecreasing estimated conditional CDFs. Together with r and m, these yield plug-in versions of F_0 and F_1. The G_B components can be similarly obtained and refined as a sequence of regressions with outcome variable equal to trimming indicator times actual outcome.

To obtain the derived conditional quantiles Q_0 and Q_1, inversion of the previously obtained conditional CDFs, F_0 and F_1, can be used. These yield the trimming indicators that are required to evaluate the conditional expectation models G_B at their respective trimming quantiles to obtain derived nuisances ψ_B. The inversion-based approach ensures algebraic compatibility between the different nuisance quantities. Thus, we use this approach in all of our simulations and applications in Section 6 as well as Appendix D and E.

4.4. Large Sample Properties

We now present additional technical assumptions and the resulting large sample properties of the efficient influence function based estimators of the causal effect bounds. Denote the true nuisances as $\eta(x) =: \eta \in \mathcal{H}$ where $\mathcal{H}$ is a convex subset of a suitably normed vector space. Denote $\mathcal{H}_n \subset \mathcal{H}$ the realization set of the estimated nuisance quantities $\hat{\eta}(x) =: \hat{\eta}$, i.e., the set containing estimated nuisances with probability $1-u_n$ where $u_n = o(1)$. All nuisances are cross-fitted according to Definition 3.2 in Chernozhukov et al. (2018), see Algorithm 1.

For the remainder, write for generic nuisance $h = h(x)$ its estimation error $\Delta h = \hat{h} - h$. Denote $\|\cdot\|_2$ as $L^2(P)$ norm and $\|\cdot\|_\infty$ as the uniform norm. By abuse of notation, if the object depends on $Z = z$ and/or $D = d$, we suppress dependence and take all norms to be uniform over these finite dimension as well, e.g.,

$$||r||_p = \sup_{z,d\in\{0,1\}} ||r(z,x)||_p = \sup_{z\in\{0,1\}} ||r(z,x)||_p = \sup_{z\in\{0,1\}} \left(\int r^p(x,z)dP(x)\right)^{1/p} \quad (4.7)$$

and equivalently for other nuisances. For a given x, we also denote $\|\cdot\|_{\infty,\mathcal{N}_x}$ as the uniform norm over a neighborhood $\mathcal{N}_x$. In particular, for a generic object $A(\cdot|x)$,

$$||\hat{A}(y|x) - A(y|x)||_{\infty,\mathcal{N}_x} = \sup_{y\in\mathcal{N}_x} |\hat{A}(y|x) - A(y|x)|. \tag{4.8}$$

We denote the supremum over these neighborhoods as

$$||\hat{A}(y|x) - A(y|x)||_{\infty,\mathcal{N}} = \sup_{x} \sup_{y\in\mathcal{N}_x} |\hat{A}(y|x) - A(y|x)|. \tag{4.9}$$

Moreover, we write shorthand

$$||\Delta G||_p = ||\Delta G_L||_{p,\mathcal{N}} + ||\Delta G_U||_{p,\mathcal{N}} \tag{4.10}$$

and $a_n \lesssim b_n$ and $a_n \lesssim_P b_n$, whenever $a_n = O(b_n)$ or $a_n = O_p(b_n)$ respectively. If not stated differently, the following assumptions are all uniformly over n.

Assumption A (Regularity, Overlap, and Learning Rates)

A.1 (Moments) The conditional potential outcome moments are bounded, i.e., for some $m > 0$,

$$\sup_{x\in\mathcal{X},d\in\{0,1\}} E[|Y(d)|^{2+m}|X = x] \lesssim 1.$$

A.2 (Eigenvalues) The variance-covariance matrix of the influence function $E[\mathbb{IF}(\beta)\mathbb{IF}(\beta)']$ has finite eigenvalues bounded away from zero.

A.3 (No Point Mass at Trimming Points) The conditional distributions are continuous at the trimming points. For $z \in \{0,1\}$, and x on the relevant support,[10]

$$\begin{aligned}
&P(Y = Q_1(p(x), x)|DS = 1, Z = z, X = x) = 0,\\
&P(Y = Q_1(1 - p(x), x)|DS = 1, Z = z, X = x) = 0,\\
&P(Y = Q_0(1 - 1/p(x), x)|(1 - D)S = 1, Z = z, X = x) = 0,\\
&P(Y = Q_0(1/p(x), x)|(1 - D)S = 1, Z = z, X = x) = 0.
\end{aligned}$$

A.4 (Bounded Mixture Outcome Density and Local Lipschitz Trimmed Means) Let $f_1(\cdot|x)$ and $f_0(\cdot|x)$ denote the respective densities of $F_1(\cdot|x)$ and $F_0(\cdot|x)$. Assume they are bounded at the trimming thresholds and the trimmed conditional means are locally Lipschitz, i.e., there exist a $C > 0$ and constants $0 < f_{\min} \le f_{\max} < \infty$, $L_G < \infty$

[10]Here and in A.4, the relevant support is $\mathcal{X}_{ac}^{+}$ for conditions involving $(Q_1(p(x), x)$ or $Q_1(1 - p(x), x)$, and $\mathcal{X}_{ac}^{-}$ for conditions involving $Q_0(1 - 1/p(x), x)$ or $Q_0(1/p(x), x)$.

such that (i) for all x on the relevant support:

$$\begin{aligned} &f_1(Q_1(p(x),x)\mid x) \in [f_{\min}, f_{\max}],\\ &f_1(Q_1(1-p(x),x)\mid x) \in [f_{\min}, f_{\max}],\\ &f_0(Q_0(1-1/p(x),x)\mid x) \in [f_{\min}, f_{\max}],\\ &f_0(Q_0(1/p(x),x)\mid x) \in [f_{\min}, f_{\max}],\end{aligned}$$

and (ii) uniformly in (z,x), for $|u| \leq C$,

$$\begin{aligned} &|G_L(Q_1(p(x),x)+u|1,z,x) - G_L(Q_1(p(x),x)|1,z,x)| \leq L_G|u|,\\ &|G_U(Q_1(1-p(x),x)+u|1,z,x) - G_U(Q_1(1-p(x),x)|1,z,x)| \leq L_G|u|,\\ &|G_L(Q_0(1-1/p(x),x)+u|0,z,x) - G_L(Q_0(1-1/p(x),x)|0,z,x)| \leq L_G|u|,\\ &|G_U(Q_0(1/p(x),x)+u|0,z,x) - G_U(Q_0(1/p(x),x)|0,z,x)| \leq L_G|u|.\end{aligned}$$

A.5 *(Margin Condition) The distribution of the positive and negative monotonicity type is well-behaved around the margin of indifference, i.e., there exist $C_M < \infty$ and $\kappa > 0$ such that*

$$P\Big(\Big| - [m(1,X) - m(0,X)] - [r(1,X) - r(0,X)]\Big| \leq t\Big) \leq C_M\, t^{\kappa} \quad \text{for all } t > 0.$$

A.6 *(Strong Overlap) There are comparable units across instrument levels and within the differently treated and selected populations. Moreover, always-selected complier probabilities are bounded away from zero on their relevant supports, i.e., there exists some $\underline{c} \in (0,1)$ such that*

$$\underline{c} < \inf_{z,x}\{r(z,x), m(z,x), e(x)\} \leq \sup_{z,x}\{r(z,x), m(z,x), e(x)\} < 1 - \underline{c}.$$

and

$$\inf_{x\in\mathcal{X}^+_{ac}} -[m(1,x) - m(0,x)] > \underline{c}, \qquad \inf_{x\in\mathcal{X}^-_{ac}} [r(1,x) - r(0,x)] > \underline{c}.$$

A.7 *(Machine Learning Bias) Let $u_n = o(1)$. For all folds, the nuisance parameters obtained via cross-fitting belong to a shrinking neighborhood $\mathcal{H}_n$ around η with probability of at least $1 - u_n$, such that, uniformly over the neighborhood, the nuisance functions are consistent*

$$\|\Delta\mu\|_2 + \|\Delta\nu\|_2 + \|\Delta e\|_2 + \|\Delta m\|_\infty + \|\Delta r\|_\infty + ||\Delta G||_{\infty,\mathcal{N}} + ||\Delta F||_{\infty,\mathcal{N}} = o(1),$$

and obey convergence rates

$$\begin{aligned}&(\|\Delta\mu\|_2 + \|\Delta\nu\|_2)\|\Delta e\|_2 + (||\Delta m||_\infty + ||\Delta r||_\infty)^{\kappa+1} + (||\Delta e||_2 + ||\Delta m||_2 + ||\Delta r||_2)\times\\ &\Bigg[||\Delta e||_\infty + ||\Delta m||_\infty + ||\Delta r||_\infty + ||\Delta G||_{\infty,\mathcal{N}} + ||\Delta F||_{\infty,\mathcal{N}}\Bigg] = o(n^{-1/2}).\end{aligned}$$

Assumption A.1 and A.2 are simple regularity conditions that rule out heavy tails and degenerate DGPs where bounds are close or equal to a point or otherwise degenerate.

Assumption A.3 rules out point masses at the trimming thresholds in the observed selected outcome distributions used to construct the reduced-form nuisance functions. This ensures that the trimming indicators are unambiguous at the cutoff and that weak and strict inequality conventions coincide at the relevant thresholds.[11]

Assumption A.4(i) is complementary to A.3 and provides primitive density regularity that guarantees local invertibility and differentiability. It controls error propagation through the quantile mapping (Bahadur-type representation). A.4(ii) adds local smoothness to the trimmed mean functions, ensuring uniform control over the nuisance functions evaluated close to the trimming points.

Assumption A.5 is a margin condition that controls the mass of observations near the boundary between positive and negative selection monotonicity types among compliers. In particular, it rules out excessive concentration of probability that would make correct classification difficult. In the case of a bounded density, A.5 holds with $\kappa = 1$, see, e.g., Audibert and Tsybakov (2007) or Heiler et al. (2024) for related assumptions and discussion. The larger κ, the less demanding the convergence requirements in A.7 for learning the conditional joint probability of being selected and in treatment/control status. It implies the necessary regularity condition $P(\mathcal{X}^0) = 0$.

Assumption A.6 assures that there are comparable units for instrument and the treatment within the selected group and a relevant share of always-selected compliers. Strong overlap is imposed to obtain a finite variance bound and avoid irregular identification (Khan and Tamer, 2010; Heiler and Kazak, 2021).

Assumption A.7 imposes the standard DML requirement that all nuisances are consistent and converge to their truth sufficiently fast for the second-order remainder of the orthogonal von Mises expansion to be $o(n^{-1/2})$, but it does so in a relatively weak form tailored to our target parameter and estimation procedure and its primitives. In particular, in contrast to much of the Lee bounds-type literature, we do not assume global uniform convergence rates for the estimated quantiles or trimmed mean functionals. Instead, it only imposes local sup-norm control of the estimated CDFs F and trimmed means G_B on neighborhoods relevant for trimming, with the regularity of quantiles and trimmed

[11] The theory can be extended to mass points as discussed in Section 2. However, in contrast to identification where mass points are harmless once one uses fractional trimming (Huber and Mellace, 2015; Kitagawa, 2021), inference can be affected as the functional may be nonregular. This problem arises only in boundary cases where the trimming probability coincides exactly with the edge of a mass point. For DGPs in which the cutoff lies in the interior of a mass point, fractional trimming yields a regular functional, so standard root-n semiparametric inference can proceed using the corresponding influence function.

means derived from these local conditions. The first product term in A.7 matches the usual L^2-rate requirement in Chernozhukov et al. (2018), while the additional terms capture the non-smooth features of monotonicity type classification and trimming indicators. This is related to rate conditions in Heiler (2024) and Semenova (2025) for generalized Lee bounds as well as Heiler et al. (2024) for more general intensive and extensive margin treatment effects, but expressed directly in terms of CDF errors rather than through separate uniform convergence assumptions on the associated quantile-based functionals. Examples for global uniform rates of nonparametric and machine learning estimation of conditional CDFs can be found in, e.g., Xie (2023) or Cattaneo et al. (2025) respectively.

We obtain the following Theorem:

Theorem 4.1 *Under Assumptions 3.1, 3.2, and A.1–A.7, the estimated bounds are jointly asymptotically normal and semiparametrically efficient, i.e.*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}\Big(0, E[\mathbb{IF}(\beta)\mathbb{IF}(\beta)']\Big).$$

Theorem 4.1 can be used directly for inference on the bounds using the usual standard normal critical values. Importantly, the assumptions imply that the identified set always has a non-empty interior in the population. Thus, Theorem 4.1 is sufficient to construct tighter confidence intervals for the effect θ_{SLATE} directly using Imbens and Manski (2004) or Stoye (2020) critical values. We use the latter in our empirical applications.

5. Always-selected Complier Profiling

θ_{SLATE} is the average treatment effect for the target population ac, the always-selected compliers. While individual members of this group are not identified and the group may differ from observed subpopulations or the overall population in both observed and unobserved characteristics, its observable features can nevertheless be characterized, much like complier profiling for the LATE.[12] This makes it possible to compare always-selected compliers with other populations of interest and thereby better assess the external validity and substantive importance of the effect bounds.

[12] See, for example, Abadie (2003), Angrist (2004), and Singh and Sun (2023) for various approaches to complier profiling.

Observable ac characteristics can be identified using the same ingredient, $\pi_{ac}(x)$, as the SLATE bounds in Section 3. Debiased estimation and inference follow analogously from our influence functions. In particular, for any integrable $g : \mathcal{X} \to \mathbb{R}$, consider target parameter

$$E[g(X)|ac] = \frac{E[g(X)\pi_{ac}(X)]}{E[\pi_{ac}(X)]}. \tag{5.1}$$

This is the average $g(x)$ in the population of always-selected compliers. Its influence function is given by

$$\mathbb{IF}(E[g(X)|ac]) = \frac{1}{\pi_{ac}}\Big((\mathbb{IF}(\pi_{ac}) + \pi_{ac})\,(g(X) - E[g(X)|ac]) \Big), \tag{5.2}$$

where nuisances and influence functions of the components can be found in Section 4.2. The corresponding estimator can be obtained by solving the empirical analogue as in Algorithm 1. It is root-n consistent and asymptotically normal analogously to Theorem 4.1. Moreover, influence function (5.2) can directly be used for statistically valid comparisons between mean covariates $g(x)$ of ac and other populations such as unconditional, treated or assigned units. We provide some specific examples in Section 6 and Appendix E.

6. Empirical Study I: Job Corps Revisited

6.1. Data and Methods

In this section, we re-evaluate the earnings effect of participating in JC, a large US federally funded training program providing free academic education, vocational training and employment assistance to disadvantaged youth. We make use of the National Job Corps Study by Mathematica Policy Research. This experiment implemented stratified randomized assignment of applicants, incorporating over 15,400 individuals between ages 16 and 24. Multiple outcomes such as earnings and job status were gathered at various points after assignment.

Our data and main variables are identical to Chen and Flores (2015).[13] In particular, we use a subset of 9,090 units (3,599 control and 5,491 treated) with non-missing work hours, earnings, and participation information. The outcome is log hourly wages which is only

[13]We would like to thank Carlos Flores for sharing the data.

observed for the employed. Assignment is given by the original randomization. Treatment is defined as eventual JC participation within the 208 weeks of the evaluation period. We additionally make use of socio-economic pre-assignment covariates including job and earnings history, education, parental background and more, matching the variables used in Lee (2009) for ITT bounds. The list of covariates along with sample summary statistics are provided in Table F.1.

The data are suitable for our method: Assignment was based on stratified randomization justifying conditional independence. Exclusion is credible as any earnings effects likely require actual training and not just assignment. Importantly, there was a significant amount of noncompliance with the randomized treatment assignment. In particular, only 73.8% of individuals assigned to the treatment group ever participated in JC. Additionally, a small fraction (4.4%) of individuals assigned to the control group also ended up participating.[14]

We evaluate the always-selected complier effect θ_{SLATE} at week 208 after assignment using (i) Chen and Flores (2015) bounds and (ii) sharp-basic bounds – both assuming strong sample selection monotonicity – as well as (iii) sharp DML bounds with covariates under weak monotonicity. Implementation of (i) follows Chen and Flores (2015). (ii) uses simple sample analogues of (2.19) and (2.20). (iii) is obtained via the procedure in Section 4.3, see Appendix F for additional details. We also provide a profiling analysis of always-selected compliers using the method from Section 5.

6.2. Results: Bounds and Shares

Table 6.1 reports the evaluation results. It delivers four main messages. First, the target population is empirically relevant: the estimated share of always-selected compliers is about 39% across specifications, so θ_{SLATE} pertains to a sizable subpopulation. Second, the data do not support imposing strong sample selection monotonicity: the estimated share in the positive sample selection region is 93.3% and strong sample selection monotonicity is rejected ($p < 0.01$).[15] Thus, Sharp-DML appears to be the most cred-

[14] Among the 4.4%, 1.2% of controls enrolled in JC before the end of the embargo, while 3.2% enrolled afterward.

[15] Negative employment effects at week 208 are economically plausible because employment responses to Job Corps are heterogeneous across applicants. Consistent with this, Semenova (2025) shows that positive conditional employment effects do not arise for all applicants even four years after random assignment and documents subgroups with significantly negative employment effects at later horizons.

Table 6.1: Bounds for the Effect of Job Corps on Log Wages

	CF	Sharp-basic	Sharp-DML
Estimates	$[-0.022,\ 0.130]$	$[0.019,\ 0.067]$	$[-0.013,\ 0.124]$
Standard Errors	—	$(0.023,\ 0.022)$	$(0.023,\ 0.022)$
95% Confidence Interval	$[-0.061,\ 0.168]$	$[-0.018,\ 0.104]$	$[-0.051,\ 0.160]$
Share of positive sample selection [a]	1 (assumed)		0.933 (0.004)
Share of *ac* [b]	0.391 (0.010)		0.393 (0.010)

Notes: $N = 9,090$. All calculations use design weights. **CF** refers to the Chen and Flores (2015) bounds under strong sample selection monotonicity without covariates using half-median-unbiased estimates with 95% confidence intervals following Chernozhukov et al. (2013). CF reports no standard errors as inference relies on the CLR projection. **Sharp-basic** refers to the sharp bounds without covariates under strong sample selection monotonicity. **Sharp-DML** refers to the sharp bounds estimated by DML under weak sample selection monotonicity. Standard errors for the two bounds are in parentheses as $(\widehat{SE}_{\text{lower}}, \widehat{SE}_{\text{upper}})$, and the 95% CI for $\theta_{SLATE} = E[Y_1 - Y_0 \mid ac]$ is calculated using Stoye (2020).

[a] Estimated share of positive sample selection $= E[\mathbb{I}^+(X)]$, the fraction of observations in the positive sample selection class (Sharp-DML only, CF and Sharp-basic assume this fraction equals 1).

[b] Estimated share of always-selected compliers π_{ac}.

ible specification. Third, under the strong monotonicity benchmark without covariates, the sharp bounds and effect confidence interval are substantially tighter than CF, by 68.4% and 46.7%, respectively. Moreover, the estimated basic bounds no longer include zero and the corresponding confidence interval rules out negative effects below -1.8%. Fourth, estimates using Sharp-DML are broadly comparable to CF despite relying only on weak sample selection monotonicity. Both rule out negative effects beyond -5.1% and -6.1% respectively. However, despite imposing less restrictive assumptions, Sharp-DML delivers tighter bounds and confidence intervals than CF by 9.8% and 7.9% respectively, highlighting the relevance of both sharpness and efficient inclusion of covariates.

The results also speak to the importance of the no-scaling result in Proposition 2.2. In particular, naively scaling basic ITT bounds $[-0.019, 0.093]$ (Lee, 2009) with the JC compliance probability of 69.4% would suggest SLATE bounds of $[-0.027, 0.134]$ which vastly exceed our Sharp-basic bounds of $[0.019, 0.067]$.

6.3. Always-selected Complier Profiling

We now conduct the *ac* profiling analysis as discussed in Section 5. Table 6.2 contains the baseline characteristics of always-selected compliers, *ac*, and those of the full sample. Demographic differences are modest but systematic: On average, *ac* individuals are less

Table 6.2: Job Corps Covariates: Always-Selected Compliers vs. Full Sample

Covariate	*ac*	Full sample	Difference
Female	0.400 (0.014)	0.443 (0.005)	−0.044 (0.013)***
Age (in yrs.) at baseline	18.45 (0.058)	18.44 (0.023)	0.013 (0.053)
Black, non-Hispanic	0.433 (0.014)	0.500 (0.005)	−0.067 (0.013)***
Hispanic	0.194 (0.010)	0.172 (0.004)	0.022 (0.009)**
Other race/ethnicity	0.067 (0.007)	0.071 (0.003)	−0.004 (0.007)
Married	0.015 (0.004)	0.022 (0.002)	−0.007 (0.004)*
Living together	0.033 (0.006)	0.041 (0.002)	−0.008 (0.005)
Separated	0.019 (0.004)	0.023 (0.002)	−0.004 (0.004)
Has children	0.162 (0.011)	0.204 (0.004)	−0.042 (0.010)***
Number of children	0.214 (0.017)	0.291 (0.007)	−0.077 (0.017)***
Education (in yrs.)	10.23 (0.042)	10.12 (0.016)	0.111 (0.039)***
Mother's education	11.54 (0.064)	11.48 (0.024)	0.061 (0.059)
Father's education	11.54 (0.060)	11.45 (0.023)	0.096 (0.054)*
Ever arrested	0.235 (0.011)	0.248 (0.005)	−0.013 (0.011)
Household income:			
[$3,000, $6,000)	0.189 (0.009)	0.208 (0.003)	−0.019 (0.008)**
[$6,000, $9,000)	0.122 (0.007)	0.116 (0.003)	0.006 (0.006)
[$9,000, $18,000)	0.258 (0.010)	0.244 (0.004)	0.014 (0.009)
≥ $18,000	0.198 (0.009)	0.179 (0.003)	0.019 (0.008)**
Personal income:			
[$3,000, $6,000)	0.130 (0.009)	0.131 (0.003)	−0.001 (0.008)
[$6,000, $9,000)	0.066 (0.006)	0.051 (0.002)	0.015 (0.005)***
≥ $9,000	0.038 (0.005)	0.033 (0.002)	0.004 (0.005)
At baseline:			
Has job	0.242 (0.011)	0.195 (0.004)	0.047 (0.010)***
Months worked, previous year	4.377 (0.119)	3.566 (0.045)	0.810 (0.109)***
Had a job, previous year	0.716 (0.013)	0.629 (0.005)	0.087 (0.012)***
Earnings, previous year	3470 (137.0)	2870 (58.00)	600.0 (111.0)***
Weekly hours, most recent job	24.35 (0.560)	21.42 (0.220)	2.930 (0.520)***
Weekly earnings, most recent job	121.8 (5.800)	107.6 (2.900)	14.10 (3.800)***

Notes: $N = 9{,}090$. The table reports estimated means for the always-selected complier (*ac*) subpopulation and for the full sample, along with their difference. Standard errors are in parentheses. All estimates use design weights and are based on the weak-monotonicity DML specification with GRF learners. Joint χ^2 tests for the null that all differences within a category are jointly zero: Race/ethnicity ($p < 0.001$); Marital status ($p = 0.084$); Household income ($p = 0.011$); Personal income ($p = 0.027$). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

likely to be female (-4.4 pp) and Black (-6.7 pp), slightly more likely to be Hispanic (+2.2 pp) and have fewer children (-4.2 pp in incidence; -0.08 in number). Own education is marginally higher (+0.11 years), with parental education and age being similar across groups. *ac* individuals are slightly less concentrated in the lowest household income bracket and more represented in the highest (1.9 pp each). The most pronounced differences arise in pre-assignment labor market outcomes: *ac* individuals are significantly more likely to be employed at baseline (+4.7 pp), more likely to have worked in the previous year (+8.7 pp), and exhibit higher labor supply and earnings, including +0.81 months worked, +$600 annual earnings, +2.9 weekly hours, and +$14 weekly earnings.

Taken together, these differences indicate that *ac* subpopulation is positively selected, in particular on baseline labor-market attachment. The stronger pre-program employment and earnings profiles suggest that *ac* may be better positioned to translate JC participation into earnings gains, but also imply more favorable counterfactual trajectories in the absence of treatment. At the same time, prior evidence on JC shows that impacts operate through channels such as increased GED, increased vocational credential attainment and reduced criminal involvement, which are not confined to the most labor-market-ready participants (Schochet et al., 2008; Flores et al., 2012). Thus, while the observable composition of the *ac* group points to potential attenuation when extrapolating our estimates for θ_{SLATE} to the full sample, baseline differences alone do not fully pin down the direction or magnitude of the extrapolation.

7. Concluding Remarks

For both identification and semiparametric inference, this paper provides a synthesis of treatment evaluation under sample selection as well as noncompliance building on Lee-type bounds and the LATE framework. Given the analytic form of our population bounds, it is relatively simple to incorporate additional restrictions, such as stochastic dominance of *ac* potential outcome distribution over that of *cc* to further tighten bounds (Zhang and Rubin, 2003; Huber and Mellace, 2015; Heiler et al., 2024). Moreover, given the simple ratio-of-linear-moment structure, our influence function components can be directly leveraged for estimation and inference on heterogeneous group-specific *ac* effect bounds by combining them with nonparametric projection methods as in Heiler (2024).

References

Abadie, A. (2003). Semiparametric Instrumental Variable Estimation of Treatment Response Models. Journal of Econometrics, 113(2):231–263.

Andersen, S. C., Beuchert, L., Heiler, P., and Nielsen, H. S. (2023). A guide to impact evaluation under sample selection and missing data: Teacher's aides and adolescent mental health.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. The Economic Journal, 114(494):C52–C83.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. Journal of the American Statistical Association, 91(434):444–455.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. Annals of Statistics, 47(2):1148–1178.

Audibert, J.-Y. and Tsybakov, A. B. (2007). Fast learning rates for plug-in classifiers. The Annals of Statistics, 35(2):608 – 633.

Bartalotti, O., Kédagni, D., and Possebom, V. (2023). Identifying marginal treatment effects in the presence of sample selection. Journal of Econometrics, 234(2):565–584.

Bickel, P. J., Klaassen, C. A., Bickel, P. J., Ritov, Y., Klaassen, J., Wellner, J. A., and Ritov, Y. (1993). Efficient and Adaptive Estimation for Semiparametric Models. Johns Hopkins University Press Baltimore.

Cattaneo, M. D., Feng, Y., and Shigida, B. (2025). Uniform estimation and inference for nonparametric partitioning-based m-estimators.

Chen, X. and Flores, C. A. (2015). Bounds on treatment effects in the presence of sample selection and noncompliance: the wage effects of job corps. Journal of Business & Economic Statistics, 33(4):523–540.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. Econometrics Journal, 21(1):C1–C68.

Chernozhukov, V., Lee, S., and Rosen, A. M. (2013). Intersection bounds: Estimation and inference. Econometrica, 81(2):667–737.

Christelis, D. and Messina, J. (2019). Partial identification of population average and quantile treatment effects in observational data under sample selection. Journal of Applied Econometrics (forthcoming).

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year*. The Quarterly Journal of Economics, 127(3):1057–1106.

Flores, C. A., Flores-Lagunes, A., Gonzalez, A., and Neumann, T. C. (2012). Estimating the effects of length of exposure to instruction in a training program: The case of job corps. Review of Economics and Statistics, 94(1):153–171.

Foresi, S. and Peracchi, F. (1995). The Conditional Distribution of Excess Returns: An Empirical Analysis. Journal of the American Statistical Association, 90(430):451–466.

Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. Biometrics, 58(1):21–29.

Frölich, M. (2007). Nonparametric IV Estimation of Local Average Treatment Effects with Covariates. Journal of Econometrics, 139(1):35–75.

Hahn, J. (1998). On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects. Econometrica, 66(2):315–331.

Heckman, J. J. and Vytlacil, E. J. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. Econometrica, 73(3):669–738.

Heiler, P. (2022). Efficient covariate balancing for the local average treatment effect. Journal of Business & Economic Statistics, 40(4):1569–1582.

Heiler, P. (2024). Heterogeneous treatment effect bounds under sample selection with an application to the effects of social media on political polarization. Journal of Econometrics, 244(1).

Heiler, P., Kaufmann, A., and Veliyev, B. (2024). Treatment evaluation at the intensive and extensive margins.

Heiler, P. and Kazak, E. (2021). Valid inference for treatment effect parameters under irregular identification and many extreme propensity scores. Journal of Econometrics, 222(2):1083–1108.

Heiler, P. and Knaus, M. C. (2026). Heterogeneity analysis with heterogeneous treatments.

Henzi, A., Ziegel, J. F., and Gneiting, T. (2021). Isotonic distributional regression. Journal of the Royal Statistical Society Series B: Statistical Methodology, 83(5):963–993.

Hines, O., Dukes, O., Diaz-Ordaz, K., and Vansteelandt, S. (2022). Demystifying statistical learning based on efficient influence functions. American Statistician, 76(3):292–304.

Horowitz, J. L. and Manski, C. F. (1995). Identification and robustness with contaminated and corrupted data. Econometrica, pages 281–302.

Horowitz, J. L. and Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. Journal of the American Statistical Association, 95(449):77–84.

Huber, M. and Mellace, G. (2015). Sharp bounds on causal effects under sample selection. Oxford Bulletin of Economics and Statistics, 77(1):129–151.

Imai, K. (2007). Identification analysis for randomized experiments with noncompliance and “truncation-by-death”.

Imai, K. (2008). Sharp bounds on the causal effects in randomized experiments with “truncation-by-death”. Statistics & Probability Letters, 78(2):144–149.

Imbens, G. W. and Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. Econometrica, 62(2):467–475.

Imbens, G. W. and Manski, C. F. (2004). Confidence intervals for partially identified parameters. Econometrica, 72(6):1845–1857.

Kennedy, E. H. (2024). Semiparametric doubly robust targeted double machine learning: a review. Handbook of statistical methods for precision medicine, pages 207–236.

Khan, S. and Tamer, E. (2010). Irregular Identification, Support Conditions, and Inverse Weight Estimation. Econometrica, 78(6):2021–2042.

Kitagawa, T. (2021). The identification region of the potential outcome distributions under instrument independence. Journal of Econometrics, 225(2):231–253. Themed Issue: Treatment Effect 1.

Klein, N. (2024). Distributional regression for data analysis. Annual Review of Statistics and Its Application, 11(Volume 11, 2024):321–346.

Lechner, M. and Melly, B. (2010). Partial identification of wage effects of training programs.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. Review of Economic Studies, 76(3):1071–1102.

Lee, Y.-Y. and Liu, C.-A. (2025). Lee bounds with a continuous treatment in sample selection.

Manski, C. F. (1997). Monotone treatment response. Econometrica, 65(6):1311–1334.

Schochet, P. Z., Burghardt, J., and McConnell, S. (2008). Does job corps work? impact findings from the national job corps study. American Economic Review, 98(5):1864–86.

Semenova, V. (2025). Generalized Lee bounds. Journal of Econometrics, 251:106055.

Singh, R. and Sun, L. (2023). Double robustness for complier parameters and a semiparametric test for complier characteristics. The Econometrics Journal, 27(1):1–20.

Stoye, J. (2020). A simple, short, but never-empty confidence interval for partially identified parameters. arXiv preprint arXiv:2010.10484.

Sun, X., Zhao, X., and Poskitt, D. S. (2024). Partially identified heterogeneous treatment effect with selection: An application to gender gaps.

Vytlacil, E. (2002). Independence, Monotonicity, and Latent Index Models: An Equivalence Result. Econometrica, 70(1):331–341.

Xie, H. (2023). Uniform convergence results for the local linear regression estimation of the conditional distribution. Statistics & Probability Letters, 201:109886.

Zhang, J. L. and Rubin, D. B. (2003). Estimation of causal effects via principal stratification when some outcomes are truncated by “death”. Journal of Educational and Behavioral Statistics, 28(4):353–368.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used OpenAI’s ChatGPT to assist with mathematical proofs, coding, formatting of tables and outputs, and spelling. The authors reviewed and edited the content as needed and take full responsibility for the content of the article.

Supplementary Appendix

A. Supplementary Material for Section 2

A.1. Proof of Proposition 2.1

The proof has two steps. Let $\beta_1 := E\left[Y_1|ac\right]$.

Step 1, we show $\beta_{L,1} \leq \beta_1 \leq \beta_{U,1}$, i.e., $\beta_{L,1}$ is the smallest feasible value and $\beta_{U,1}$ is the largest feasible value for β_{11} that is consistent with the observed data and maintained assumptions. It suffices to show $\beta_1 \geq \beta_{L,1}$, and the other part $\beta_1 \leq \beta_{U,1}$ can be proved similarly. We have shown that under Assumptions 2.1–2.2 and assuming $\pi_{ac} > 0$, $F_{Y_1|ac\cup cc}\left(y\right)$ is uniquely identified. In particular,

$$F_{Y_1|ac\cup cc}\left(y\right) = \frac{E\left[DS \cdot \mathbb{1}\left\{Y \leq y\right\}|Z=1\right] - E\left[DS \cdot \mathbb{1}\left\{Y \leq y\right\}|Z=0\right]}{E\left[DS|Z=1\right] - E\left[DS|Z=0\right]}.$$

Further,

$$F_{Y_1|ac\cup cc}\left(y\right) = pF_{Y_1|ac}\left(y\right) + \left(1-p\right)F_{Y_1|cc}\left(y\right)$$

where p is uniquely identified:

$$p = \frac{E\left[\left(D-1\right)S|Z=1\right] - E\left[\left(D-1\right)S|Z=0\right]}{E\left[DS|Z=1\right] - E\left[DS|Z=0\right]}.$$

If $\pi_{cc} = 0$, then $p = 1$ and $ac \cup cc = ac$, so $E[Y_1 \mid ac]$ is point-identified and $\beta_{L,1} = \beta_{U,1} = E[Y_1 \mid ac]$. Hence, in what follows, suppose $0 < p < 1$.

$\beta_{L,1}$ assumes

$$F_{Y_1|ac}\left(y\right) = F^{dh}_{Y_1|ac}\left(y\right) := \begin{cases} \frac{1}{p}F_{Y_1|ac\cup cc}\left(y\right) & \text{if } y < Q_1\left(p\right) \\ 1 & \text{if } y \geq Q_1\left(p\right) \end{cases}$$

and hence

$$\beta_{L,1} = \int_{-\infty}^{\infty} yF^{dh}_{Y_1|ac}\left(dy\right) = \frac{1}{p}\int_{-\infty}^{Q_1(p)} yF_{Y_1|ac\cup cc}\left(dy\right).$$

By Horowitz and Manski (1995), Corollary 4.1, $\beta_1 \geq \beta_{L,1}$. Further, $\beta_{L,1}$ is smallest feasible value that is consistent with the observed data because both $F_{Y_1|ac\cup cc}\left(y\right)$ and p are uniquely determined by the observed data under the maintained assumptions.

Step 2, we show that every point in $[\beta_{L,1}, \beta_{U,1}]$ is compatible with our assumptions and the observed data and hence cannot be ruled out, which implies that $[\beta_{L,1}, \beta_{U,1}]$ is contained in any other valid bounds that impose the same assumptions.

Note that $\beta_{L,1} \leq E\left[Y_1|ac \cup cc\right] \leq E\left[Y_1|ac \cup cc, Y_1 > Q_1\left(p\right)\right]$, where both conditional mean functions are uniquely identified, as the distribution function $F_{Y_1|ac\cup cc}$ and the quantile level p are uniquely identified.

For any point δ between $\beta_{L,1}$ and $E\left[Y_1|ac \cup cc\right]$, there exists a value $\lambda \in [0,1]$ such that

$$\delta = \lambda\beta_{L,1} + \left(1-\lambda\right)E\left[Y_1|ac \cup cc, Y_1 > Q_1\left(p\right)\right]$$

Let

$$h_l\left(y\right) = \begin{cases} \frac{1}{p}F_{Y_1|ac\cup cc}\left(y\right) & \text{if } y < Q_1\left(p\right) \\ 1 & \text{if } y \geq Q_1\left(p\right) \end{cases}$$

and

$$h_u\left(y\right) = \begin{cases} 0 & \text{if } y < Q_1\left(p\right) \\ \frac{1}{1-p}\left(F_{Y_1|ac\cup cc}\left(y\right) - p\right) & \text{if } y \geq Q_1\left(p\right) \end{cases}$$

We can construct distribution functions for $Y_1|ac$ and $Y_1|cc$, respectively, as

$$\begin{aligned} F_{Y_1|ac}\left(y\right) &= \lambda h_l\left(y\right) + \left(1-\lambda\right)h_u\left(y\right), \\ F_{Y_1|cc}\left(y\right) &= \left(\frac{p}{1-p} - \frac{p}{1-p}\lambda\right)h_l\left(y\right) + \left(1 - \frac{p}{1-p}\left(1-\lambda\right)\right)h_u\left(y\right). \end{aligned}$$

Note

$$E[Y_1 \mid ac \cup cc] = p\,\beta_{L,1} + (1-p)\,E[Y_1 \mid ac \cup cc,\ Y_1 > Q_1(p)],$$

and $\delta \in [\beta_{L,1},\ E[Y_1 \mid ac \cup cc]]$, so the corresponding λ must satisfy $\lambda \in [p, 1]$. Hence,

$$0 \leq \frac{p}{1-p}(1-\lambda) \leq 1.$$

The coefficients used below in the definition of $F_{Y_1|cc}$ are then nonnegative and sum to one, meaning that the constructed $F_{Y_1|ac}$ and $F_{Y_1|cc}$ are valid distribution functions. By construction,

$$\begin{aligned} pF_{Y_1|ac}\left(y\right) + \left(1-p\right)F_{Y_1|cc}\left(y\right) &= ph_l\left(y\right) + \left(1-p\right)h_u\left(y\right) \\ &= F_{Y_1|ac\cup cc}\left(y\right). \end{aligned}$$

i.e., the mixture of these two distribution functions replicates the uniquely identified distribution function $F_{Y_1|ac\cup cc}(y)$and hence they are compatible with our assumptions and the observed data. Further by construction,

$$\begin{aligned}
E[Y_1|ac] = \int_{-\infty}^{+\infty} y F_{Y_1|ac}(dy) &= \int_{-\infty}^{+\infty} y \{\lambda h_l(dy) + (1-\lambda) h_u(dy)\} \\
&= \lambda \int_{-\infty}^{+\infty} y h_l(dy) + (1-\lambda) \int_{-\infty}^{+\infty} y h_u(dy) \\
&= \lambda \beta_{L,1} + (1-\lambda) E[Y_1|ac \cup cc, Y_1 > Q_1(p)] \\
&= \delta.
\end{aligned}$$

That is, δ is a feasible point that is compatible with our assumptions and the observed data.

A symmetric argument can be made about any point δ in between $E[Y_1|ac \cup cc]$ and $\beta_{U,1}$. Therefore, any point within the interval $[\beta_{L,1}, \beta_{U,1}]$ is feasible and compatible with our assumptions and observed data.

Together, the above two results, 1) $\beta_{L,1} \leq \beta_1 \leq \beta_{U,1}$, and 2) any point within the interval $[\beta_{L,1}, \beta_{U,1}]$ is feasible and compatible with our assumptions and observed data, suggest $\beta_{L,1}$ ($\beta_{U,1}$) is sharp, i.e., $\beta_{L,1}$ ($\beta_{U,1}$) is the largest (smallest) lower (upper) bound for $\beta_1 := E[Y_1|ac]$ that is consistent with our assumptions and observed data.

A.2. Supplementary Material for Section 2.2.1

A.2.1. Lee (2009) ITT bounds and SLATE bounds

Below we formalize the relationship between the Lee (2009) bounds and our SLATE bounds. Recall the complier types defined in Section 2.1, and define in addition

$$an := \{S_0 = 1,\, D_0 = D_1 = 0\} \quad \text{(always-selected never-takers)},$$

$$aa := \{S_1 = 1,\, D_0 = D_1 = 1\} \quad \text{(always-selected always-takers)}.$$

We also allow for types that are never selected,

$$\begin{aligned}
nc &:= \{S_0 = S_1 = 0,\, D_0 = 0, D_1 = 1\}, \\
na &:= \{S_1 = 0,\, D_0 = D_1 = 1\}, \\
nn &:= \{S_0 = 0,\, D_0 = D_1 = 0\},
\end{aligned}$$

which will play no role in the Lee-type selected distributions because $S = 1$ never occurs for them. Let $\pi_t := P(\text{type } t)$ denote the population share of type t.

Lemma A.1 (Decomposition of selected distributions) *Suppose Assumptions 2.1–2.2 hold and the relevant latent types exist. Then the distribution of Y among the selected units satisfies*

$$F_{Y|S=1,Z=1}(y) = \omega\, F_{Y_0|an}(y) + (1-\omega)\, F_{Y_1|ac\cup cc\cup aa}(y), \tag{A.1}$$

$$\tag{A.2}$$

$$F_{Y|S=1,Z=0}(y) = \delta\, F_{Y_0|ac\cup an}(y) + (1-\delta)\, F_{Y_1|aa}(y), \tag{A.3}$$

where

$$\omega := \frac{\pi_{an}}{\pi_{ac}+\pi_{cc}+\pi_{aa}+\pi_{an}}, \qquad \delta := \frac{\pi_{ac}+\pi_{an}}{\pi_{ac}+\pi_{an}+\pi_{aa}}.$$

Proof.

$$\begin{aligned} F_{Y|S=1,Z=1}\left(y\right) &= F_{Y|(1-D)SZ=1}\left(y\right) E\left[(1-D)SZ|SZ=1\right] \\ &+ F_{Y|DSZ=1}\left(y\right) E\left[DSZ|SZ=1\right] \\ &= F_{Y_0|S_0=1,D_1=0}\left(y\right) \frac{E\left[S_0\left(1-D_1\right)\right]}{E\left[S_1D_1\right]+E\left[S_0\left(1-D_1\right)\right]} \\ &+ F_{Y_1|S_1=1,D_1=1}\left(y\right) \frac{E\left[S_1D_1\right]}{E\left[S_1D_1\right]+E\left[S_0\left(1-D_1\right)\right]} \\ &= F_{Y_0|an}\left(y\right)\omega + F_{Y_1|ac\cup cc\cup aa}\left(y\right)\left(1-\omega\right), \end{aligned}$$

where $\omega{:=}\frac{\pi_{an}}{\pi_{ac}+\pi_{cc}+\pi_{aa}+\pi_{an}}$.

Similarly,

$$\begin{aligned} F_{Y|S=1,Z=0}\left(y\right) &= F_{Y|S=1,D=0,Z=0}\left(y\right) E\left[(1-D)S\left(1-Z\right)|S\left(1-Z\right)=1\right] \\ &+ F_{Y|S=1,D=1,Z=0}\left(y\right) E\left[DS\left(1-Z\right)|S\left(1-Z\right)=1\right] \\ &= F_{Y_0|S_0=1,D_0=0}\left(y\right) \frac{E\left[S_0\left(1-D_0\right)\right]}{E\left[S_1D_0\right]+E\left[S_0\left(1-D_0\right)\right]} \\ &+ F_{Y_1|S_1=1,D_0=1}\left(y\right) \frac{E\left[S_1D_0\right]}{E\left[S_1D_0\right]+E\left[S_0\left(1-D_0\right)\right]} \\ &= F_{Y_0|ac\cup an}\left(y\right)\delta + F_{Y_1|aa}\left(y\right)\left(1-\delta\right), \end{aligned}$$

where $\delta{:=}\frac{\pi_{ac}+\pi_{an}}{\pi_{ac}+\pi_{aa}+\pi_{an}}$

We can now make precise the relationship between Lee (2009) lower bound and our lower bound for SLATE. Let β_L denote our lower bound for SLATE,

$$\beta_L = E\left[Y_1|Y_1 \leq Q_1(p),\, ac\cup cc\right] - E[Y_0|ac],$$

with $p := \frac{\pi_{ac}}{\pi_{ac}+\pi_{cc}}$ is the fraction of always-selected compliers within the complier mixture and $Q_1\left(u\right) := \inf\left\{y\in\mathcal{Y} : F_{Y_1|ac\cup cc}\left(y\right)\geq u\right\}$

Let β_L^{Lee} denote the Lee (2009) lower bound for the ITT effect of Z on Y among the selected, as in (2.26). Recall $p^{Lee} := \frac{E[S|Z=0]}{E[S|Z=1]} = \frac{\pi_{ac}+\pi_{aa}+\pi_{an}}{\pi_{ac}+\pi_{cc}+\pi_{aa}+\pi_{an}}$, and $Q_1^{Lee}\left(u\right) :=$

$\inf\left\{y \in \mathcal{Y} : F_{Y_0|an}(y)\,\omega + F_{Y_1|ac\cup cc\cup aa}(y)\,(1-\omega) \geq u\right\}$. For notational convenience, let the trimming threshold be $q := Q_1^{Lee}\left(p^{Lee}\right)$. Note that evaluating (A.2) at $y = q$ yields

$$p^{Lee} = \omega F_{Y_0|an}(q) + (1-\omega)\, F_{Y_1|ac\cup cc\cup aa}(q)\,.$$

Combining Lemma A.1 with the definition of β_L^{Lee}, one can write the first term in β_L^{Lee} as

$$\begin{aligned}\beta_{L,1}^{Lee} &:= E\big[Y \mid Y \leq Q_1^{Lee}(p^{Lee}),\, S=1,\, Z=1\big] \\ &= w(q)\, E[Y_0 \mid Y_0 \leq q,\, an] + \big(1 - w(q)\big)\, E[Y_1 \mid Y_1 \leq q,\, ac \cup cc \cup aa]\,,\end{aligned}$$

where

$$w(q) = \frac{\omega F_{Y_0|an}(q)}{p^{Lee}}$$

and the second term in β_L^{Lee} as

$$\begin{aligned}\beta_0^{Lee} &:= E\left[Y|S=1, Z=0\right] \\ &= E\left[Y_0|ac \cup an\right]\delta + E\left[Y_1|aa\right](1-\delta)\,.\end{aligned}$$

So together,

$$\begin{aligned}\beta_L^{Lee} = &\,\{\text{weighted average of } Y_0 \text{ for } an \text{ and } Y_1 \text{ for } ac, cc, aa\} \\ &- \{\text{weighted average of } Y_0 \text{ for } ac, an \text{ and } Y_1 \text{ for } aa\},\end{aligned}$$

where the exact weights depend on $(\pi_{ac}, \pi_{cc}, \pi_{an}, \pi_{aa})$, the trimming fraction p^{Lee}, and the component CDFs entering $F_{Y|S=1,Z=1}$ evaluated at the trimming threshold q.

A.2.2. Proof of Proposition 2.2

Proof. By Lemma A.1, the first term in β_L^{Lee} is a trimmed mean of Y over the mixture $\{ac \cup cc \cup aa \cup an\}$ in the selected treated group, and the second term is an untrimmed mean over a different mixture of types $\{ac, an, aa\}$ in the selected control group. Both mixtures involve always-takers and/or never-takers whenever $\pi_{an} + \pi_{aa} > 0$.

By contrast, β_L depends only on complier types (ac and cc): its Y_1 component is a trimmed mean of Y_1 restricted to $\{ac \cup cc\}$, and its Y_0 component is the mean of Y_0 for always-selected compliers (ac). Unless $\pi_{an} = \pi_{aa} = 0$, the distributions of Y_0 and Y_1 entering β_L^{Lee} necessarily differ from those entering β_L because of the additional type-specific components.

Holding the primitive probabilities fixed, one can vary the distribution of Y_0 for an

and/or the distribution of Y_1 for aa while leaving the distributions for ac and cc unchanged; this changes β_L^{Lee} through the additional mixture components, and in the treated term also through the trimming threshold q and the effective an stratum weight $w(q)$, while leaving β_L unchanged.

Therefore, for generic joint distributions of (Y_0, Y_1) across types, β_L^{Lee} and β_L are distinct functionals, and they cannot be linked uniformly over DGPs by a single multiplicative function depending only on the type probabilities.

A.3. Supplementary Material for Section 2.2.2

In this appendix we formalize the comparison between our bounds and those of Chen and Flores (2015). The strict comparisons are under the maintained Section 2 regularity for the trimmed-mean representation.

A.3.1. Identification of $F_{Y_1|ac\cup cc}$

We first show how the distribution of Y_1 for the complier mixture (ac and cc) can be recovered from observable selected distributions and identified type shares.

Lemma A.2 *Suppose Assumptions 2.1–2.2 hold and $\pi_{ac} > 0$. If $\pi_{aa} > 0$, then*

$$F_{Y_1|ac\cup cc}(y) = \frac{(\pi_{ac} + \pi_{cc} + \pi_{aa})F_{Y_1|ac\cup cc\cup aa}(y) - \pi_{aa}F_{Y_1|aa}(y)}{\pi_{ac} + \pi_{cc}}.$$

Equivalently, in terms of observable distributions,

$$F_{Y_1|ac\cup cc}(y) = \frac{(\pi_{ac} + \pi_{cc} + \pi_{aa})F_{Y|DSZ=1}(y) - \pi_{aa}F_{Y|DS(1-Z)=1}(y)}{\pi_{ac} + \pi_{cc}}, \tag{A.4}$$

where

$$F_{Y|DSZ=1}(y) = F_{Y_1|ac\cup cc\cup aa}(y), \qquad F_{Y|DS(1-Z)=1}(y) = F_{Y_1|aa}(y).$$

If $\pi_{aa} = 0$, then (A.4) *reduces to $F_{Y_1|ac\cup cc}(y) = F_{Y|DSZ=1}(y)$.*

Proof. Among units with $DSZ = 1$, the only types that can appear are ac, cc, and aa, all with $S_1 = 1$ and $D_1 = 1$. Thus when $\pi_{aa} > 0$,

$$F_{Y|DSZ=1}(y) = F_{Y_1|ac\cup cc\cup aa}(y).$$

Otherwise,

$$F_{Y|DSZ=1}(y) = F_{Y_1|ac\cup cc}(y).$$

Further when $\pi_{aa} > 0$, among units with $DS(1-Z) = 1$ the only possible type is aa, or those with $S_1 = 1$ and $D_0 = 1$, so

$$F_{Y|DS(1-Z)=1}(y) = F_{Y_1|aa}(y).$$

Writing $F_{Y_1|ac\cup cc\cup aa}$ as a mixture of $F_{Y_1|ac\cup cc}$ and $F_{Y_1|aa}$, we have

$$F_{Y_1|ac\cup cc\cup aa}(y) = \frac{\pi_{ac}+\pi_{cc}}{\pi_{ac}+\pi_{cc}+\pi_{aa}} F_{Y_1|ac\cup cc}(y) + \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}} F_{Y_1|aa}(y),$$

and solving for $F_{Y_1|ac\cup cc}(y)$ yields (A.4).

A.3.2. Dominance over CF bounds: Preliminaries

For clarity, observe the following equivalences of conditional distributions:

$$\big(Y_1 \mid S_1 = 1,\ D_1 > D_0\big) \equiv \big(Y_1 \mid ac \cup cc\big), \qquad \big(Y \mid DSZ = 1\big) \equiv \big(Y_1 \mid ac \cup cc \cup aa\big),$$
$$\big(Y \mid DS(1-Z) = 1\big) \equiv \big(Y_1 \mid aa\big).$$

As in the main text, let $\beta_{L,1}$ denote the first component (pertaining to Y_1) of our lower bound (i.e. excluding β_0),

$$\beta_{L,1} = E\left[Y_1 | Y_1 \leq Q_1(p), ac \cup cc\right], \qquad p := \frac{\pi_{ac}}{\pi_{ac}+\pi_{cc}}.$$

Let $\beta^{mix}_{L,1}$ and $\beta^{adj}_{L,1}$ be the first components of the CF basic and alternative lower bounds obtained by trimming $Y|DSZ = 1$ or equivalently $Y_1|ac \cup cc \cup aa$ at fractions p^{CF}_1 and p^{CF}_2:

$$\beta^{mix}_{L,1} = E\left[Y_1 | Y_1 \leq Q^{CF}_1(p^{CF}_1), ac \cup cc \cup aa\right], \qquad p^{CF}_1 := \frac{\pi_{ac}}{\pi_{ac}+\pi_{cc}+\pi_{aa}},$$

$$\beta^{adj}_{L,1} = \left(1 + \frac{\pi_{aa}}{\pi_{ac}}\right) E\left[Y_1 | Y_1 \leq Q^{CF}_1(p^{CF}_2), ac \cup cc \cup aa\right] - \frac{\pi_{aa}}{\pi_{ac}} E\left[Y_1 | aa\right], \qquad p^{CF}_2 := \frac{\pi_{ac}+\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}}.$$

The CF lower bound uses

$$\beta^{CF}_{L,1} := \max\{\beta^{mix}_{L,1}, \beta^{adj}_{L,1}\}.$$

The formal proof of Proposition 2.3 is provided below.

Proof. We sketch the argument for the lower bounds; the upper bounds are treated analogously by symmetry.

First of all, if $\pi_{aa} = 0$, i.e., $Pr\left(DS\left(1-Z\right)=1\right)=0$, then $F_{Y_1|ac\cup cc\cup aa}\left(y\right)$ reduces to $F_{Y_1|ac\cup cc}\left(y\right)$ and the trimming threshold $p_1^{CF} = p_2^{CF} = p$, so $\beta_{L,1} = \beta_{L,1}^{CF}$ and $\beta_{U,1} = \beta_{U,1}^{CF}$. In the following, we assume $\pi_{aa} > 0$.

$\beta_{L,1}$ is obtained by assuming that ac individuals are at the bottom $p{:=}\frac{\pi_{ac}}{\pi_{ac}+\pi_{cc}}$ fraction of the conditional distribution $Y_1|ac \cup cc$, i.e., $\beta_{L,1}$ assumes

$$F_{Y_1|ac}\left(y\right) = F_{Y_1|ac}^{dh}\left(y\right) := \begin{cases} \frac{1}{p} F_{Y_1|ac\cup cc}\left(y\right) & \text{if } y < Q_1\left(p\right) \\ 1 & \text{if } y \geq Q_1\left(p\right) \end{cases} \tag{A.5}$$

where $Q_1\left(p\right) := F_{Y_1|ac\cup cc}^{-1}\left(p\right)$. Following Equation (A.4),

$$F_{Y_1|ac\cup cc}\left(y\right) = \left(1 + \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}}\right) F_{Y_1|ac\cup cc\cup aa}(y) - \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}} F_{Y_1|aa}(y). \tag{A.6}$$

Plugging in Equation (A.6) into (A.5) shows that $\beta_{L,1}$ assumes

$$F_{Y_1|ac}\left(y\right) = F_{Y_1|ac}^{dh}\left(y\right) = \begin{cases} \frac{1}{p_1^{CF}} F_{Y_1|ac\cup cc\cup aa}(y) - \frac{\pi_{aa}}{\pi_{ac}} F_{Y_1|aa}(y) & \text{if } y < Q_1\left(p\right) \\ 1 & \text{if } y \geq Q_1\left(p\right) \end{cases} \tag{A.7}$$

In contrast, $\beta_{L,1}^{mix}$ is obtained by assuming that ac individuals are at the bottom $p_1^{CF}{:=}\frac{\pi_{ac}}{\pi_{ac}+\pi_{cc}+\pi_{aa}}$ fraction of the $Y_1|ac \cup cc \cup aa$ distribution (the CF restriction 1), i.e.,

$$F_{Y_1|ac}\left(y\right) = F_{Y_1|ac}^{mix}\left(y\right) := \begin{cases} \frac{1}{p_1^{CF}} F_{Y_1|ac\cup cc\cup aa}(y) & \text{if } y < Q_1^{cf}\left(p_1^{CF}\right) \\ 1 & \text{if } y \geq Q_1^{cf}\left(p_1^{CF}\right) \end{cases}, \tag{A.8}$$

where $Q_1^{cf}\left(p_1^{CF}\right) := F_{Y_1|ac\cup cc\cup aa}^{-1}(p_1^{CF})$.

Putting Equations (A.7) and (A.8) together and noticing $Q_1^{cf}\left(p_1^{CF}\right) \leq Q_1\left(p\right)$, we have

$$F_{Y_1|ac}^{mix}(y) - F_{Y_1|ac}^{dh}(y) = \begin{cases} \frac{\pi_{aa}}{\pi_{ac}} F_{Y_1|aa}(y), & y < Q_1^{cf}\left(p_1^{CF}\right), \\ 1 - \frac{1}{p} F_{Y_1|ac\cup cc}(y) > 0, & Q_1^{cf}\left(p_1^{CF}\right) \leq y < Q_1(p), \\ 0, & y \geq Q_1(p). \end{cases} \tag{A.9}$$

To show $Q_1^{cf}\left(p_1^{CF}\right) \leq Q_1\left(p\right)$, evaluating Equation (A.6) at $y = Q_1\left(p\right)$ and rearranging yields

$$F_{Y_1|ac\cup cc\cup aa}(Q_1(p)) = p_1^{CF} + \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}} F_{Y_1|aa}(Q_1(p)) \tag{A.10}$$
$$\geq p_1^{CF} = F_{Y_1|ac\cup cc\cup aa}(Q_1^{cf}(p_1^{CF})).$$

$Q_1^{cf}\left(p_1^{CF}\right) \leq Q_1(p)$ follows from that $F_{Y_1|ac\cup cc\cup aa}(\cdot)$ is strictly increasing.

Note that under the maintained regularity, Equation (A.6) implies the following equivalence:

$$F_{Y_1|aa}(Q_1(p)) > 0 \iff Q_1^{cf}(p_1^{CF}) < Q_1(p) \iff F_{Y_1|aa}(Q_1^{cf}(p_1^{CF})) > 0.$$

Equation (A.9) implies $F^{mix}_{Y_1|ac}(y) \geq F^{dh}_{Y_1|ac}(y)$ in general, i.e., the conditional distribution $Y_1|ac$ assumed in $\beta_{L,1}$ weakly stochastically dominates that assumed in $\beta^{mix}_{L,1}$. Furthermore, $F^{mix}_{Y_1|ac}(y) > F^{dh}_{Y_1|ac}(y)$ iff $\pi_{aa} > 0$ and $F_{Y_1|aa}(Q_1(p)) > 0$. The mean respects stochastic dominance, so

$$\beta_{L,1} \geq \beta^{mix}_{L,1}$$

in general, as well as

$$\beta_{L,1} > \beta^{mix}_{L,1}, \quad \text{iff} \quad \pi_{aa} > 0 \quad \text{and} \quad F_{Y_1|aa}(Q_1(p)) > 0.$$

Now consider $\beta^{adj}_{L,1}$, which is based on the CF restriction 2. It assumes that ac and aa are at the bottom $p_2^{CF} := \frac{\pi_{ac}+\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}}$ fraction of the conditional distribution $Y_1|ac\cup cc\cup aa$, i.e.

$$F_{Y_1|ac\cup aa}(y) = F^{adj}_{Y_1|ac\cup aa}(y) := \begin{cases} \frac{1}{p_2^{CF}} F_{Y_1|ac\cup cc\cup aa}(y) & \text{if } y < Q_1^{cf}\left(p_2^{CF}\right) \\ 1 & \text{if } y \geq Q_1^{cf}\left(p_2^{CF}\right) \end{cases}. \tag{A.11}$$

In contrast, we have shown that $\beta_{L,1}$ assumes

$$F_{Y_1|ac}(y) = F^{dh}_{Y_1|ac}(y) := \begin{cases} \frac{1}{p_1^{CF}} F_{Y_1|ac\cup cc\cup aa}(y) - \frac{\pi_{aa}}{\pi_{ac}} F_{Y_1|aa}(y) & \text{if } y < Q_1(p) \\ 1 & \text{if } y \geq Q_1(p) \end{cases},$$

which implies $\beta_{L,1}$ assumes, for $F_{Y_1|ac\cup aa}(y)$,

$$F_{Y_1|ac\cup aa}(y) = F^{dh}_{Y_1|ac\cup aa}(y) := \frac{\pi_{ac}}{\pi_{ac}+\pi_{aa}} F^{dh}_{Y_1|ac}(y) + \frac{\pi_{aa}}{\pi_{ac}+\pi_{aa}} F_{Y_1|aa}(y)$$
$$= \begin{cases} \frac{1}{p_2^{CF}} F_{Y_1|ac\cup cc\cup aa}(y) & \text{if } y < Q_1(p) \\ 1 - \frac{\pi_{aa}}{\pi_{ac}+\pi_{aa}}\left(1 - F_{Y_1|aa}(y)\right) & \text{if } y \geq Q_1(p) \end{cases}. \tag{A.12}$$

Putting together Equations (A.11) and (A.12) and noticing $Q_1^{cf}\left(p_2^{CF}\right) \geq Q_1(p)$ yield

$$
\begin{aligned}
&F^{alt}_{Y_1|ac\cup aa}(y) - F^{dh}_{Y_1|ac\cup aa}(y) \\
&= \begin{cases} 0, & y < Q_1(p), \\ \dfrac{\pi_{ac}+\pi_{cc}}{\pi_{ac}+\pi_{aa}} \left(F_{Y_1|ac\cup cc}(y) - p\right) \geq 0, & Q_1(p) \leq y < Q_1^{cf}\left(p_2^{CF}\right), \\ \dfrac{\pi_{aa}}{\pi_{ac}+\pi_{aa}} \left(1 - F_{Y_1|aa}(y)\right), & y \geq Q_1^{cf}\left(p_2^{CF}\right). \end{cases}
\end{aligned} \tag{A.13}
$$

where the middle branch uses (A.6).

To show $Q_1^{cf}\left(p_2^{CF}\right) \geq Q_1(p)$, plugging into Equation (A.10) $p_1^{CF} = p_2^{CF} - \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}}$ and rearranging yields

$$
\begin{aligned}
F_{Y_1|ac\cup cc\cup aa}(Q_1(p)) &= p_2^{CF} - \frac{\pi_{aa}}{\pi_{ac}+\pi_{cc}+\pi_{aa}} \left(1 - F_{Y_1|aa}(Q_1(p))\right) \\
&\leq p_2^{CF} \\
&= F_{Y_1|ac\cup cc\cup aa}(Q_1^{cf}\left(p_2^{CF}\right)).
\end{aligned}
$$

$Q_1^{cf}\left(p_2^{CF}\right) \geq Q_1(p)$ follows, since $F_{Y_1|ac\cup cc\cup aa}(\cdot)$ is strictly increasing. Under the maintained regularity, Equation (A.6) implies $F_{Y_1|aa}(Q_1^{cf}(p_2^{CF})) < 1 \iff Q_1^{cf}(p_2^{CF}) > Q_1(p) \iff F_{Y_1|aa}(Q_1(p)) < 1$.

Equation (A.13) implies $F^{dh}_{Y_1|ac\cup aa}(y) \leq F^{adj}_{Y_1|ac\cup aa}(y)$ in general, i.e., the conditional distribution $Y_1|ac\cup aa$ assumed in $\beta_{L,1}$ weakly stochastically dominates that assumed in $\beta^{adj}_{L,1}$. Furthermore, $F^{dh}_{Y_1|ac\cup aa}(y) < F^{alt}_{Y_1|ac\cup aa}(y)$ iff $\pi_{aa} > 0$ and $F_{Y_1|aa}(Q_1(p)) < 1$. The mean respects stochastic dominance, so in general

$$
\frac{\pi_{ac}}{\pi_{ac}+\pi_{aa}}\beta_{L,1} + \frac{\pi_{aa}}{\pi_{ac}+\pi_{aa}} E\left[Y_1|aa\right] \geq E\left[Y|DSZ=1, Y \leq Q_1^{cf}\left(p_2^{CF}\right)\right] \tag{A.14}
$$

where the left-hand side is $E\left[Y_1|ac\cup aa\right]$ assuming $F_{Y_1|ac\cup aa}(y) = F^{dh}_{Y_1|ac\cup aa}(y)$, while the right-hand side is $E\left[Y_1|ac\cup aa\right]$ assuming $F_{Y_1|ac\cup aa}(y) = F^{adj}_{Y_1|ac\cup aa}(y)$.

Rearranging Equation (A.14), we have, in general,

$$
\begin{aligned}
\beta_{L,1} &\geq \left(1 + \frac{\pi_{aa}}{\pi_{ac}}\right) E\left[Y|DSZ=1, Y \leq Q_1^{cf}\left(p_2^{CF}\right)\right] - \frac{\pi_{aa}}{\pi_{ac}} E\left[Y_1|aa\right] \\
&=: \beta^{adj}_{L,1},
\end{aligned}
$$

and further,

$$
\beta_{L,1} > \beta^{adj}_{L,1}, \quad \text{iff} \quad \pi_{aa} > 0 \quad \text{and} \quad F_{Y_1|aa}(Q_1(p)) < 1.
$$

The above shows (I) $\beta_{L,1} \geq \beta_{L,1}^{mix}$ in general and $\beta_{L,1} > \beta_{L,1}^{mix}$ iff $\pi_{aa} > 0$ and $F_{Y_1|aa}(Q_1(p)) > 0$ and (II) $\beta_{L,1} \geq \beta_{L,1}^{adj}$ in general and $\beta_{L,1} > \beta_{L,1}^{adj}$ iff $\pi_{aa} > 0$ and $F_{Y_1|aa}(Q_1(p)) < 1$. Together, they imply

$$\beta_{L,1} \geq \beta_{L,1}^{CF} := \max\{\beta_{L,1}^{mix}, \beta_{L,1}^{adj}\} \quad \text{in general,}$$

and

$$\beta_{L,1} > \beta_{L,1}^{CF}, \quad \text{iff} \quad \pi_{aa} > 0 \quad \text{and} \quad 0 < F_{Y_1|aa}(Q_1(p)) < 1.$$

The intuition is as follows. By construction, $\beta_{L,1}$ is the sharp worst-case lower bound for $E[Y_1 \mid ac]$ given the identified complier mixture $F_{Y_1|ac\cup cc}$ and the fact that ac makes up the fraction p of that mixture. The worst case is attained by placing all ac individuals below all cc individuals in $Y_1 \mid ac \cup cc$.

Viewed inside the larger mixture $Y_1 \mid ac \cup cc \cup aa$, this configuration implies only that the ac mass must be contained in the bottom

$$1 - \frac{\pi_{cc}}{\pi_{ac} + \pi_{cc} + \pi_{aa}} = p_2^{CF}$$

fraction of the distribution. Its exact location within that region is otherwise unrestricted. The CF basic bound corresponds to the most unfavorable placement within this class, namely the one in which ac occupies exactly the bottom

$$p_1^{CF} = \frac{\pi_{ac}}{\pi_{ac} + \pi_{cc} + \pi_{aa}}$$

fraction. In that limiting case, $\beta_{L,1} = \beta_{L,1}^{mix}$. Any less adverse placement shifts some ac mass upward within the bottom p_2^{CF} fraction, which raises $E[Y_1 \mid ac]$ and hence $\beta_{L,1}$, while leaving $\beta_{L,1}^{mix}$ unchanged. Therefore,

$$\beta_{L,1} \geq \beta_{L,1}^{mix}.$$

For $\beta_{L,1}^{adj}$, CF Restriction 2 requires that ac and aa together occupy the bottom p_2^{CF} fraction of $Y_1 \mid ac \cup cc \cup aa$, and then uses the identified mean $E[Y_1 \mid aa]$ to back out a lower bound on $E[Y_1 \mid ac]$. Under our sharp worst-case configuration, ac still lies entirely below cc, whereas aa may be located anywhere in the mixture. If all aa mass also lies in the bottom p_2^{CF} fraction, then $\beta_{L,1} = \beta_{L,1}^{adj}$. If some aa mass is shifted to higher quantiles, the lower bound implied by CF Restriction 2 becomes smaller, while

$\beta_{L,1}$ remains unchanged. Hence,

$$\beta_{L,1} \geq \beta_{L,1}^{adj}.$$

Combining the two comparisons yields

$$\beta_{L,1} \geq \beta_{L,1}^{CF} := \max\{\beta_{L,1}^{mix}, \beta_{L,1}^{adj}\}.$$

The upper-bound comparison follows by the same argument applied to the upper tails.

B. Supplementary Material for Section 3

B.1. Moment representations with covariates

Under Assumption 3.1, the conditional versions of Lemma 2.1 imply,

$$\lambda_1(x) := P(S_1 = 1, D_1 > D_0 | X = x) = E[W(Z, X)\, DS | X = x], \tag{B.1}$$

$$\lambda_0(x) := P(S_0 = 1, D_1 > D_0 | X = x) = E[W(Z, X)\, (D - 1)S | X = x], \tag{B.2}$$

and, for any y,

$$F_1\,(y|x) := F_{Y_1|S_1=1,D_1>D_0,X=x}(y) = \frac{E[W(Z, X)\, DS \cdot \mathbb{1}\{Y \le y\} | X = x]}{E[W(Z, X) DS | X = x]}, \tag{B.3}$$

$$F_0\,(y|x) := F_{Y_0|S_0=1,D_1>D_0,X=x}(y) = \frac{E[W(Z, X)\, (D - 1)S \cdot \mathbb{1}\{Y \le y\} | X = x]}{E[W(Z, X)(D - 1)S | X = x]}. \tag{B.4}$$

where the inverse probability weight $W(Z, X) := \frac{Z}{e(X)} - \frac{1-Z}{1-e(X)}$ with $e(x) := P(Z = 1 | X = x)$. The conditional quantiles $Q_d(u, x)$, $d = 0, 1$, used in (3.9)-(3.12) are defined as the generalized inverses of these CDFs.

Using these identities and the same arguments as in Section 2.1, one can verify the numerator moment function for β_L. For $x \in \mathcal{X}^+$, it is

$$\begin{aligned}
\beta_L^+\,(x)\,\pi_{ac}\,(x) &= \beta_L^+\,(x)\,P\,(S_0 = 1, D_1 > D_0 | X = x) \\
&= E\,[S_1 Y_1 \cdot \mathbb{1}\,\{Y_1 \le Q_1\,(p\,(x)\,, x)\} \cdot \mathbb{1}\,\{D_1 > D_0\}\,|X = x] \\
&- E\,[S_0 Y_0 \cdot \mathbb{1}\,\{D_1 > D_0\}\,|X = x] \\
&= E\,[W\,(Z, X)\,(DSY \cdot \mathbb{1}\,\{Y \le Q_1\,(p\,(x)\,, x)\} + (1 - D)\,SY)\,|X = x]
\end{aligned}$$

where the first equality follows from Assumption 3.2, the second from the definition of $\beta_L^+(x)$ and the law of iterated expectations, and the last equality follows from the conditional version of Lemma 2.1, applied with $g(y) = y\,\mathbb{1}\{y \le Q_1(p(x), x)\}$ for the treated term and with D replaced by $(D - 1)$ for the untreated term.

Similarly for $x \in \mathcal{X}^-$, the numerator moment function is

$$\begin{aligned}
\beta_L^-\,(x)\,\pi_{ac}\,(x) &= \beta_L^-\,(x)\,P\,(S_1 = 1, D_1 > D_0 | X = x) \\
&= E\,[W\,(Z, X)\,(DSY + (1 - D)\,SY \cdot \mathbb{1}\,\{Y \ge Q_0\,(1 - 1/p\,(x)\,, x)\})\,|X = x]\,.
\end{aligned}$$

Analogously, one can derive the numerator moment function for β_U. For $x \in \mathcal{X}^+$, it is

$$\begin{aligned}
\beta_U^+\,(x)\,\pi_{ac}\,(x) &= \beta_U^+\,(x)\,P\,(S_0 = 1, D_1 > D_0 | X = x) \\
&= E\,[S_1 Y_1 \cdot \mathbb{1}\,\{Y_1 \ge Q_1\,(1 - p\,(x)\,, x)\} \cdot \mathbb{1}\,\{D_1 > D_0\}\,|X = x] \\
&- E\,[S_0 Y_0 \cdot \mathbb{1}\,\{D_1 > D_0\}\,|X = x] \\
&= E\,[W\,(Z, X)\,(DSY \cdot \mathbb{1}\,\{Y \ge Q_1\,(1 - p\,(x)\,, x)\} + (1 - D)\,SY)\,|X = x]\,.
\end{aligned}$$

Similarly, for $x \in \mathcal{X}^{-}$, it is

$$\begin{aligned}\beta_U^{-}(x)\,\pi_{ac}(x) &= \beta_U^{-}(x)\,P\left(S_1 = 1, D_1 > D_0 | X = x\right)\\ &= E\left[W\left(Z, X\right)\left(DSY + \left(1-D\right)SY \cdot \mathbb{1}\left\{Y \leq Q_0\left(1/p\left(x\right), x\right)\right\}\right)|X = x\right].\end{aligned}$$

|

C. Supplementary Material for Section 4

C.1. Derivation of Influence Functions

C.1.1. Unconditional Case with Strong Sample Selection Monotonicity

Basic Decomposition

Under strong sample selection monotonicity and no covariates for both bounds we can decompose[16]

$$\beta_B = \beta_{B,1} + \beta_0,$$

where

$$\beta_{B,1} = -\frac{1}{(m_1 - m_0)}[r_1\psi_{B,1} - r_0\psi_{B,0}],$$
$$\beta_0 = -\frac{1}{(m_1 - m_0)}[\mu_1 - \mu_0],$$

where we denote shorthand

$$m_z = E[(1-D)S|Z=z],$$
$$r_z = E[DS|Z=z],$$
$$\mu_z = E[(1-D)SY|Z=z],$$
$$\psi_{L,z} = E[Y\mathbb{1}(Y \leq Q_1(p))|DS=1, Z=z],$$
$$\psi_{U,z} = E[Y\mathbb{1}(Y > Q_1(1-p))|DS=1, Z=z].$$

Now consider a parametric submodel indexed by $t \in (0,1]$. From the definition and the product rule it follows that

$$-[m_1 - m_0]\frac{\partial}{\partial t}\beta_{B,1,t}\Big|_{t=0} = \frac{\partial}{\partial t}r_{1,t}\Big|_{t=0}\psi_{B,1} - \frac{\partial}{\partial t}r_{0,t}\Big|_{t=0}\psi_{B,0} + r_1\frac{\partial}{\partial t}\psi_{B,1,t}\Big|_{t=0} - r_0\frac{\partial}{\partial t}\psi_{B,0,t}\Big|_{t=0} + \beta_{B,1}\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}]\Big|_{t=0}$$

and

$$-[m_1 - m_0]\frac{\partial}{\partial t}\beta_{0,t}\Big|_{t=0} = \frac{\partial}{\partial t}[\mu_{1,t} - \mu_{0,t}]\Big|_{t=0} + \beta_0\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}]\Big|_{t=0}.$$

We omit denoting the evaluation of the derivative at $t=0$ in what follows whenever it does not cause confusion.

[16]Note that β_0 here is used different from Section 2.1, i.e., it is the sign-adjusted control mean (+ instead of $-$ in the definition of β_B). This is for notational simpilcity only.

Quantile Auxiliary Results

Differentiating the result of Lemma 2.1 for $g(Y) = \mathbb{1}(Y \leq y)$ with respect to y yields density function

$$f_{Y_1|S_1=1,D_1>D_0}(y) = \frac{r_1 f_1(y) - r_0 f_0(y)}{r_1 - r_0},$$

where $f_z(y) = f_{Y|DS=1,Z=z}(y)$. Using the definition of the quantile as well as Leibniz' rule for differentiation then yields

$$\begin{aligned}
p &= \int^{Q_1(p)} f_{Y_1|S_1=1,D_1>D_0}(y)dy \\
\Rightarrow -[m_1 - m_0] &= p[r_1 - r_0] \\
&= r_1 \int^{Q_1(p)} f_1(y)dy - r_0 \int^{Q_1(p)} f_0(y)dy \\
\Rightarrow -\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] &= \frac{\partial}{\partial t} r_{1,t} F_1(Q_1(p)) - \frac{\partial}{\partial t} r_{0,t} F_0(Q_1(p)) \\
&\quad + [r_1 f_1(Q_1(p)) - r_0 f_0(Q_1(p))] \frac{\partial}{\partial t} Q_{1,t}(p_t) \\
&\quad + r_1 \int^{Q_1(p)} \frac{\partial}{\partial t} f_{1,t}(y)dy - r_0 \int^{Q_1(p)} \frac{\partial}{\partial t} f_{0,t}(y)dy \\
\Rightarrow [r_1 f_1(Q_1(p)) - r_0 f_0(Q_1(p))] &\frac{\partial}{\partial t} Q_{1,t}(p_t) \\
&= -\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] - \left[\frac{\partial}{\partial t} r_{1,t} F_1(Q_1(p)) - \frac{\partial}{\partial t} r_{0,t} F_0(Q_1(p))\right] \\
&\quad - \left[r_1 \int^{Q_1(p)} \frac{\partial}{\partial t} f_{1,t}(y)dy - r_0 \int^{Q_1(p)} \frac{\partial}{\partial t} f_{0,t}(y)dy\right].
\end{aligned}$$

Equivalently

$$\begin{aligned}
p &= \int_{Q_1(1-p)} f_{Y_1|S_1=1,D_1>D_0}(y)dy \\
\Rightarrow -[m_1 - m_0] &= p[r_1 - r_0] \\
&= r_1 \int_{Q_1(1-p)} f_1(y)dy - r_0 \int_{Q_1(1-p)} f_0(y)dy \\
\Rightarrow -\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] &= \frac{\partial}{\partial t} r_{1,t}[1 - F_1(Q_1(1-p))] - \frac{\partial}{\partial t} r_{0,t}[1 - F_0(Q_1(1-p))] \\
&\quad - [r_1 f_1(Q_1(1-p)) - r_0 f_0(Q_1(1-p))] \frac{\partial}{\partial t} Q_{1,t}(1-p_t) \\
&\quad + r_1 \int_{Q_1(1-p)} \frac{\partial}{\partial t} f_{1,t}(y)dy - r_0 \int_{Q_1(1-p)} \frac{\partial}{\partial t} f_{0,t}(y)dy \\
\Rightarrow -[r_1 f_1(Q_1(p)) - r_0 f_0(Q_1(p))] &\frac{\partial}{\partial t} Q_{1,t}(1-p_t) \\
&= -\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] - \left[\frac{\partial}{\partial t} r_{1,t}[1 - F_1(Q_1(1-p))] - \frac{\partial}{\partial t} r_{0,t}[1 - F_0(Q_1(1-p))]\right] \\
&\quad - \left[r_1 \int_{Q_1(1-p)} \frac{\partial}{\partial t} f_{1,t}(y)dy - r_0 \int_{Q_1(1-p)} \frac{\partial}{\partial t} f_{0,t}(y)dy\right].
\end{aligned}$$

Trimmed Mean Auxiliary Result

Applying Leibniz' rule to the trimmed means yields

$$\begin{aligned}
\frac{\partial}{\partial t}\psi_{L,z,t} &= \frac{\partial}{\partial t}\int^{Q_{1,t}(p_t)} y f_{z,t}(y)dy \\
&= Q_1(p) f_z(Q_1(p))\frac{\partial}{\partial t}Q_{1,t}(p_t) + \int^{Q_1(p)} y\frac{\partial}{\partial t}f_{z,t}(y)dy \\
\frac{\partial}{\partial t}\psi_{U,z,t} &= \frac{\partial}{\partial t}\int_{Q_{1,t}(1-p_t)} y f_{z,t}(y)dy \\
&= -Q_1(1-p) f_z(Q_1(1-p))\frac{\partial}{\partial t}Q_{1,t}(1-p_t) + \int_{Q_1(1-p)} y\frac{\partial}{\partial t}f_{z,t}(y)dy,
\end{aligned}$$

which implies that

$$\begin{aligned}
r_1\frac{\partial}{\partial t}\psi_{L,1,t} - r_0\frac{\partial}{\partial t}\psi_{L,0,t} &= Q_1(p)[r_1 f_1(Q_1(p)) - r_0 f_0(Q_1(p))]\frac{\partial}{\partial t}Q_{1,t}(p_t) \\
&\quad + \int^{Q_1(p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy \\
r_1\frac{\partial}{\partial t}\psi_{U,1,t} - r_0\frac{\partial}{\partial t}\psi_{U,0,t} &= -Q_1(1-p)[r_1 f_1(Q_1(1-p)) - r_0 f_0(Q_1(1-p))]\frac{\partial}{\partial t}Q_{1,t}(1-p_t) \\
&\quad + \int_{Q_1(1-p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy.
\end{aligned}$$

Scaled β_B Pathwise Derivative

Plugging in both auxiliary results thus yields

$$\begin{aligned}
-[m_1 - m_0]\frac{\partial}{\partial t}\beta_{L,1,t} &= \frac{\partial}{\partial t}r_{1,t}\psi_{L,1} - \frac{\partial}{\partial t}r_{0,t}\psi_{L,0} + \beta_{L,1}\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] \\
&- Q_1(p)\bigg(\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] + \frac{\partial}{\partial t}r_{1,t}F_1(Q_1(p)) - \frac{\partial}{\partial t}r_{0,t}F_0(Q_1(p)) \\
&\quad + \int^{Q_1(p)}\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}\right]dy\bigg) + \int^{Q_1(p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}\right]dy,
\end{aligned}$$

as well as

$$\begin{aligned}
-[m_1 - m_0]\frac{\partial}{\partial t}\beta_{U,1,t} &= \frac{\partial}{\partial t}r_{1,t}\psi_{U,1} - \frac{\partial}{\partial t}r_{0,t}\psi_{U,0} + \beta_{U,1}\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] \\
&- Q_1(1-p)\bigg(\frac{\partial}{\partial t}[m_{1,t} - m_{0,t}] + \frac{\partial}{\partial t}r_{1,t}[1 - F_1(Q_1(1-p))] - \frac{\partial}{\partial t}r_{0,t}[1 - F_0(Q_1(1-p))] \\
&\quad + \int_{Q_1(1-p)}\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}\right]dy\bigg) + \int_{Q_1(1-p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y) - r_0\frac{\partial}{\partial t}f_{0,t}\right]dy.
\end{aligned}$$

Influence Function Primitive Components

In the following, we provide the influence function of the primitives via the pathwise derivative of the parameter $\beta_B(P)$ in a fully nonparametric model satisfying Assumptions A.1–A.7 without covariates X. We use the previous derivations and standard influence functions for all unrestricted conditional mean primitives, see, e.g., Kennedy (2024).

$$\begin{aligned}
\frac{\partial}{\partial t}\mu_{z,t} &= \frac{\mathbb{1}(Z=z)}{P(Z=z)}((1-D)SY-\mu(z)),\\
\Rightarrow \frac{\partial}{\partial t}[\mu_{1,t}-\mu_{0,t}] &= W(Z)((1-D)SY-\mu(Z)),
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial t}m_{z,t} &= \frac{\mathbb{1}(Z=z)}{P(Z=z)}((1-D)S-m(z)),\\
\Rightarrow \frac{\partial}{\partial t}[m_{1,t}-m_{0,t}] &= W(Z)((1-D)S-m(Z)),\\
\frac{\partial}{\partial t}r_{z,t} &= \frac{\mathbb{1}(Z=z)}{P(Z=z)}(DS-r(z)),\\
\Rightarrow \frac{\partial}{\partial t}r_{1,t}\psi_{B,1}-\frac{\partial}{\partial t}r_{0,t}\psi_{B,0} &= W(Z)(DS-r(Z))\psi_B(Z),\\
\frac{\partial}{\partial t}r_{1,t}F_1(Q_1(p))-\frac{\partial}{\partial t}r_{0,t}F_0(Q_1(p)) &= W(Z)(DS-r(Z))F(Q_1(p)|1,Z),\\
\frac{\partial}{\partial t}r_{1,t}[1-F_1(Q_1(1-p))]-\frac{\partial}{\partial t}r_{0,t}[1-F_0(Q_1(1-p))] &= W(Z)(DS-r(Z))[1-F(Q_1(1-p)|1,Z)],
\end{aligned}$$

where $\psi_B(Z)=Z\psi_{B,1}+(1-Z)\psi_{B,0}$ and $F(u|Z)=ZF_1(u)+(1-Z)F_0(u)$. Now consider the conditional probability $f_z(y)=f(y|DS=1,Z=z)$. By standard arguments, its influence function is given by

$$\begin{aligned}
\frac{\partial}{\partial t}f_{z,t}(y) &= \frac{1(Z=z)DS}{P(DS=1|Z=z)P(Z=z)}(\mathbb{1}(Y=y)-f_z(y))\\
\Rightarrow r_z\int^{Q_1(p)}\frac{\partial}{\partial t}f_{z,t}(y)dy &= \frac{1(Z=z)DS}{P(Z=z)}\left(\mathbb{1}(Y\leq Q_1(p))-\int^{Q_1(p)}f_z(y)dy\right)\\
&= \frac{1(Z=z)DS}{P(Z=z)}\left(\mathbb{1}(Y\leq Q_1(p))-F_z(Q_1(p))\right).
\end{aligned}$$

Thus we have that

$$\int^{Q_1(p)}\left[r_1\frac{\partial}{\partial t}f_{1,t}(y)-r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy = W(Z)DS\Big(\mathbb{1}(Y\leq Q_1(p))-F(Q_1(p)|1,Z)\Big)$$

and, by similar derivations,

$$\int^{Q_1(p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y)-r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy = W(Z)DS\Big(Y\mathbb{1}(Y\le Q_1(p))-\psi_L(Z)\Big).$$

The upper bound components are found analogously as

$$\int_{Q_1(1-p)}\left[r_1\frac{\partial}{\partial t}f_{1,t}(y)-r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy = W(Z)DS\Big(\mathbb{1}(Y>Q_1(1-p))-[1-F(Q_1(1-p)|1,Z)]\Big)$$

and

$$\int_{Q_1(1-p)} y\left[r_1\frac{\partial}{\partial t}f_{1,t}(y)-r_0\frac{\partial}{\partial t}f_{0,t}(y)\right]dy = W(Z)DS\Big(Y\mathbb{1}(Y>Q_1(1-p))-\psi_U(Z)\Big).$$

Influence Function for Combined Rescaled β_0 and β_1

Plugging in the previous result into the rescaled pathwise derivative of β_0 yields

$$-[m_1-m_0]\frac{\partial}{\partial t}\beta_{0,t} = W(Z)((1-D)SY-\mu(Z))+\beta_0 W(Z)((1-D)S-m(Z)).$$

For the lower bound we obtain

$$\begin{aligned}-[m_1-m_0]\frac{\partial}{\partial t}\beta_{L,1} &= W(Z)(DS-r(Z))\psi_L(Z)+\beta_{L,1}W(Z)((1-D)S-m(Z))\\ &\quad - Q_1(p)W(Z)\Big[((1-D)S-m(Z))+(DS-r(Z))F(Q_1(p)|1,Z)\\ &\qquad + DS(\mathbb{1}(Y\le Q_1(p))-F(Q_1(p)|1,Z))\Big]\\ &\quad + W(Z)DS(Y\mathbb{1}(Y\le Q_1(p))-\psi_L(Z)),\end{aligned}$$

while for the upper bound we have

$$\begin{aligned}-[m_1-m_0]\frac{\partial}{\partial t}\beta_{U,1} &= W(Z)(DS-r(Z))\psi_U(Z)+\beta_{U,1}W(Z)((1-D)S-m(Z))\\ &\quad - Q_1(1-p)W(Z)\Big[((1-D)S-m(Z))+(DS-r(Z))[1-F(Q_1(1-p)|1,Z)]\\ &\qquad + DS(\mathbb{1}(Y>Q_1(1-p))-[1-F(Q_1(1-p)|1,Z)])\Big]\\ &\quad + W(Z)DS(Y\mathbb{1}(Y>Q_1(1-p))-\psi_U(Z)).\end{aligned}$$

C.1.2. Conditional Case with Strong Sample Selection Monotonicity

We now add covariates to the problem. Recall that, for any bound β_B, by definition,

$$\begin{aligned}
\beta_B &= \frac{\int \beta_B(x)\pi_{ac}(x)dP(x)}{\int \pi_{ac}(x)dP(x)} \\
&= \frac{1}{\int \pi_{ac}(x)dP(x)}\left[\int \beta_{B,1}(x)\pi_{ac}(x)dP(x) + \int \beta_0(x)\pi_{ac}(x)dP(x)\right] \\
&\equiv \frac{N_{B,1}+N_0}{\pi_{ac}}.
\end{aligned}$$

For simplicity and without loss of generality, we now treat X as taking values in a countable set. In general, the tangent space of the model is the closure of the linear span of scores that are constant on finitely many measurable subsets of the support of X. Coarsening X by such a finite partition produces a discrete analogue X^{disc} for which the calculation below applies verbatim. As the partition becomes finer, the functional $\beta_B^{disc}(P)$ converge to $\beta_B(P)$, and the corresponding pathwise derivatives converge as well, with Assumptions A.3–A.5 ensuring the passing limits through the derivative. Since scores are dense in the tangent space and the derivative operator is continuous, the influence functions obtained under the discrete-X approximation will also be valid for general X. In particular, we consider point-mass perturbations in the direction $O = (YS, S, D, Z, X)$ using submodels of the form $dP_t = (1-t)\,dP_0 + t\,\delta_o$, where $P_0 \in \mathcal{P}$ denotes the true law and $\mathcal{P}$ the nonparametric model. By standard results for conditional means (e.g., Bickel et al. (1993), Ch. 3), these perturbations span the tangent space. For any functional parameter, $\mathbb{IF}$ is then influence function operator that maps from the functional to its Riesz representer under the point mass perturbation submodel. We obtain

$$\mathbb{IF}(\beta_B) = \frac{1}{\pi_{ac}}\left[\mathbb{IF}(N_{B,1}) + \mathbb{IF}(N_0) - \beta_B \mathbb{IF}(\pi_{ac})\right].$$

To derive the components first note the following auxiliary results

$$\begin{aligned}
\mathbb{IF}(\pi_{ac}(x)) &= -\mathbb{IF}(m(1,x) - m(0,x)) \\
\mathbb{IF}(m(z,x)) &= \frac{\mathbb{1}(Z=z, X=x)}{P(Z=z|X=x)P(X=x)}((1-D)S - m(z,x)) \\
\Rightarrow \mathbb{IF}(\pi_{ac}(x))P(X=x) &= -\mathbb{1}(X=x)W(Z,X)((1-D)S - m(Z,X)),
\end{aligned}$$

where $W(z,x) = z/P(Z=1|X=x) - (1-z)/(1-P(Z=1|X=x))$. The applicability of the operator to the conditional case follows from dominated convergence. Now note that

$$
\begin{aligned}
\mathbb{IF}(\pi_{ac}) &= \mathbb{IF}\left(\sum_x \pi_{ac}(x)P(X=x)\right) \\
&= \sum_x \mathbb{IF}(\pi_{ac}(x))P(X=x) + \pi_{ac}(X) - \pi_{ac} \\
&= -\Big[W(Z,X)((1-D)S - m(Z,X)) + [m(1,X) - m(0,X)]\Big] - \pi_{ac}.
\end{aligned}
$$

Next note that, for any B and $j = 0, 1$, we obtain decomposition

$$
\begin{aligned}
&\mathbb{IF}(N_j) = \mathbb{IF}\left(\sum_x \beta_{B,j}(x)\pi_{ac}(x)P(X=x)\right) \\
&= \sum_x \mathbb{IF}(\beta_{B,j}(x))\pi_{ac}(x)P(X=x) + \sum_x \beta_{B,j}(x)\mathbb{IF}(\pi_{ac}(x))P(X=x) + \sum_x \beta_{B,j}(x)\pi_{ac}(x)\mathbb{1}(X=x) - N_j
\end{aligned}
$$

and equivalently for $\beta_0 = \beta_{B,0}$. Now note that, as $\pi_{ac}(x) = -[m(1,x) - m(0,x)]$, we have that

$$
\pi_{ac}(x)\mathbb{IF}(\beta_{B,1}(x))P(X=x) = \mathbb{1}(X=x) \times \Big[-[m(1,x) - m(0,x)] \times \text{“IF of } \beta_{B,1} \text{ conditional on } x\text{”}\Big].
$$

For example

$$
\begin{aligned}
&\pi_{ac}(x)\mathbb{IF}(\beta_0(x))P(X=x) \\
&\quad = \mathbb{1}(X=x)\Big[W(Z,X)((1-D)SY - \mu(Z,X)) + \beta_0(X)W(Z,X)((1-D)S - m(Z,X))\Big]
\end{aligned}
$$

and analogously for $\beta_{B,1}$. Now also note that

$$
\beta_{B,j}(x)\mathbb{IF}(\pi_{ac}(x))P(X=x) = -\mathbb{1}(X=x)\beta_{B,j}(X)W(Z,X)((1-D)S - m(Z,X)).
$$

Lastly, by definition, we can write

$$
\begin{aligned}
-[m(1,x) - m(0,x)]\beta_0(x) &= \mu(1,x) - \mu(0,x), \\
-[m(1,x) - m(0,x)]\beta_{B,1}(x) &= \psi_B(1,x)r(1,x) - \psi_B(0,x)r(0,x).
\end{aligned}
$$

Thus, combining expressions yields

$$
\begin{aligned}
\mathbb{IF}(N_0) &= W(Z,X)((1-D)SY - \mu(Z,X)) + \beta_0(X)W(Z,X)((1-D)S - m(Z,X)) \\
&\quad - \beta_0(X)W(Z,X)((1-D)S - m(Z,X)) - \beta_0(X)(m(1,X) - m(0,X)) - N_0 \\
&= W(Z,X)((1-D)SY - \mu(Z,X)) + [\mu(1,X) - \mu(0,X)] - N_0
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{IF}(N_{L,1}) = {} & W(Z,X)(DS - r(Z,X))\psi_L(Z,X) \\
& - Q_1(p(X),X)W(Z,X)\Big[((1-D)S - m(Z,X)) \\
& \quad + (DS - r(Z,X))F(Q_1(p(X),X)|1,Z,X) \\
& \qquad + DS(\mathbb{1}(Y \leq Q_1(p(X),X)) - F(Q_1(p(X),X)|1,Z,X))\Big] \\
& + W(Z,X)DS[Y\mathbb{1}(Y \leq Q_1(p(X),X)) - \psi_L(Z,X)] \\
& + \psi_L(1,X)r(1,X) - \psi_L(0,X)r(0,X) - N_{L,1},
\end{aligned}
$$

as well as

$$
\begin{aligned}
\mathbb{IF}(N_{U,1}) = {} & W(Z,X)(DS - r(Z,X))\psi_U(Z,X) \\
& - Q_1(1-p(X),X)W(Z,X)\Big[((1-D)S - m(Z,X)) \\
& \quad + (DS - r(Z,X))[1 - F(Q_1(1-p(X),X)|1,Z,X)] \\
& \qquad + DS(\mathbb{1}(Y > Q_1(1-p(X),X)) - [1 - F(Q_1(1-p(X),X)|1,Z,X)])\Big] \\
& + W(Z,X)DS[Y\mathbb{1}(Y > Q_1(1-p(X),X)) - \psi_U(Z,X)] \\
& + \psi_U(1,X)r(1,X) - \psi_U(0,X)r(0,X) - N_{U,1},
\end{aligned}
$$

or simplified

$$
\begin{aligned}
\mathbb{IF}(N_{L,1}) = {} & -Q_1(p(X),X)W(Z,X)\Big[((1-D)S - m(Z,X)) \\
& \quad + (DS - r(Z,X))F(Q_1(p(X),X)|1,Z,X) \\
& \qquad + DS(\mathbb{1}(Y \leq Q_1(p(X),X)) - F(Q_1(p(X),X)|1,Z,X))\Big] \\
& + W(Z,X)[DSY\mathbb{1}(Y \leq Q_1(p(X),X)) - \psi_L(Z,X)r(Z,X)] \\
& + \psi_L(1,X)r(1,X) - \psi_L(0,X)r(0,X) - N_{L,1}, \\
\mathbb{IF}(N_{U,1}) = {} & -Q_1(1-p(X),X)W(Z,X)\Big[((1-D)S - m(Z,X)) \\
& \quad + (DS - r(Z,X))[1 - F(Q_1(1-p(X),X)|1,Z,X)] \\
& \qquad + DS(\mathbb{1}(Y > Q_1(1-p(X),X)) - [1 - F(Q_1(1-p(X),X)|1,Z,X)])\Big] \\
& + W(Z,X)[DSY\mathbb{1}(Y > Q_1(1-p(X),X)) - \psi_U(Z,X)r(Z,X)] \\
& + \psi_U(1,X)r(1,X) - \psi_U(0,X)r(0,X) - N_{U,1}.
\end{aligned}
$$

C.1.3. Conditional Case with Weak Sample Selection Monotonicity

Recall, that by linearity, we have that the influence function of the bounds can be decomposed as influence functions of the following main elements

$$
\mathbb{IF}(\beta_B) = \frac{1}{\pi_{ac}}\left[\mathbb{IF}(N_0^+) + \mathbb{IF}(N_0^-) + \mathbb{IF}(N_{B,1}^+) + \mathbb{IF}(N_{B,1}^-) - \beta_B \mathbb{IF}(\pi_{ac})\right].
$$

Next note that by Assumption A.5, we have that all the components are pathwise differentiable with the classification error vanishing under the integral (Heiler et al., 2024). Using the analogous derivations as for the positive monotonicity case, we obtain

$$\begin{aligned}\mathbb{IF}(\pi_{ac}) = &-\mathbb{1}^+\left[W(Z,X)((1-D)S - m(Z,X)) + m(1,X) - m(0,X)\right] \\ &+ \mathbb{1}^-\left[W(Z,X)(DS - r(Z,X)) + r(1,X) - r(0,X)\right] - \pi_{ac}\end{aligned}$$

and as before, up to a multiplier $\mathbb{1}^+$-based modification,

$$\begin{aligned}\mathbb{IF}(N_0^+) + N_0^+ &= \mathbb{1}^+(\mathbb{IF}(N_0) + N_0), \\ \mathbb{IF}(N_{B,1}^+) + N_{B,1}^+ &= \mathbb{1}^+(\mathbb{IF}(N_{B,1}) + N_{B,1}),\end{aligned}$$

and fully new components

$$\begin{aligned}
\mathbb{IF}(N_0^-) &= \mathbb{1}^-\left\{W(Z,X)(DSY - \nu(Z,X)) + \nu(1,X) - \nu(0,X)\right\} - N_0^-, \\
\mathbb{IF}(N_{L,1}^-) &= \mathbb{1}^-\Bigg\{ - Q_0(1 - \frac{1}{p(X)}, X) W(Z,X) \Bigg[(DS - r(Z,X)) \\
&\quad + ((1-D)S - m(Z,X)) \left(1 - F(Q_0(1 - \frac{1}{p(X)}, X)|0,Z,X)\right) \\
&\quad\quad + (1-D)S\left(\mathbb{1}(Y \geq Q_0(1 - \frac{1}{p(X)}, X)) - \left[1 - F(Q_0(1-\frac{1}{p(X)}, X)|0,Z,X\right]\right)\Bigg] \\
&\quad + W(Z,X)\left((1-D)SY\mathbb{1}(Y \geq Q_0(1 - \frac{1}{p(X)}, X)) - \psi_L^-(Z,X)m(Z,X)\right) \\
&\quad + \psi_L^-(1,x)m(1,x) - \psi_L^-(0,x)m(0,x)\Bigg\} - N_{L,1}^-, \\
\mathbb{IF}(N_{U,1}^-) &= \mathbb{1}^-\Bigg\{ - Q_0(\frac{1}{p(X)}, X) W(Z,X) \Bigg[(DS - r(Z,X)) \\
&\quad + ((1-D)S - m(Z,X)) \left(F(Q_0(\frac{1}{p(X)}, X)|0,Z,X)\right) \\
&\quad\quad + (1-D)S\left(\mathbb{1}(Y \leq Q_0(\frac{1}{p(X)}, X)) - F(Q_0(\frac{1}{p(X)}, X)|0,Z,X)\right)\Bigg] \\
&\quad + W(Z,X)\left((1-D)SY\mathbb{1}(Y \leq Q_0(\frac{1}{p(X)}, X)) - \psi_U^-(Z,X)m(Z,X)\right) \\
&\quad + \psi_U^-(1,x)m(1,x) - \psi_U^-(0,x)m(0,x)\Bigg\} - N_{U,1}^-.
\end{aligned}$$

C.2. Bounding the Machine Learning Bias Remainder

C.2.1. Preliminaries, Structure of Influence Functions and Remainder

For the following, at any x, let $F_1(\cdot|x)$ denote the true mixture CDF and $\hat{F}_1(\cdot|x)$ its estimator. For the proof, we denote shorthand $Q_1(x) = Q_1(p(x), x)$ and $\hat{Q}(x) = \hat{Q}(\hat{p}(x), x)$ obtained by inversion the mixture CDF. In particular, they are given by the generalized left-continuous inverses at the threshold $p(x)$ and $\hat{p}(x)$ respectively.

$$Q_1(x) = \inf\{y \in \mathbb{R} : F_1(y|x) \geq p(x)\}, \quad \hat{Q}_1(x) = \inf\{y \in \mathbb{R} : \hat{F}_1(y|x) \geq \hat{p}(x)\}.$$

We use the local neighborhood $\mathcal{N}_x := [Q_1(x) - \xi, Q_1(x) + \xi]$. For generic $\tilde{\eta}$, let

$$\Psi(\tilde{\eta}) := E\big[\mathbb{IF}(\beta_B, \tilde{\eta})\big]$$

denote the population moment of the rescaled/stabilized efficient influence function, and let η denote the true nuisance vector. The Neyman orthogonality of the EIF implies that for every admissible direction h,

$$D\Psi[\eta](h) := \lim_{t \to 0} \frac{\Psi(\eta + th) - \Psi(\eta)}{t} = E\big[\partial_\eta \mathbb{IF}(\beta_B, \eta)[h]\big] = 0.$$

For directions h_1, h_2 in the nuisance space define second directional derivative

$$D^2\Psi[\eta](h_1, h_2) := \lim_{t \to 0} \frac{D\Psi[\eta + th_2](h_1) - D\Psi[\eta](h_1)}{t},$$

whenever the limit exists. Under (A.3), (A.4), (A.5), and (A.6), $D^2\Psi[\eta]$ exists in a neighborhood of η. For $\Delta\eta = \hat{\eta} - \eta$, define the path $\eta_t = \eta + t\Delta\eta$, $t \in [0, 1]$. A second-order expansion with integral remainder yields

$$\Psi(\eta + \Delta\eta) - \Psi(\eta) = \int_0^1 (1 - t)\, D^2\Psi[\eta_t](\Delta\eta, \Delta\eta)\, dt,$$

since $D\Psi[\eta] = 0$. The second-order term is

$$R_n(\beta_B) := \Psi(\hat{\eta}) - \Psi(\eta).$$

The cross-fitted stabilized estimating equation uses the EIF

$$\mathbb{IF}(\beta_B) = \frac{\mathbb{IF}(N_0^+) + \mathbb{IF}(N_0^-) + \mathbb{IF}(N_{B,1}^+) + \mathbb{IF}(N_{B,1}^-) - \beta_B\, \mathbb{IF}(\pi_{ac})}{E[\pi_{ac}(X)]}.$$

We now denote shorthand in what follows $\mathbb{IF}\beta_B = \mathbb{IF}(\beta_B, \eta)$ and $\hat{\mathbb{IF}}(\beta_B) = \mathbb{IF}(\beta_B, \hat{\eta})$ being the EIF with estimated nuisances. By construction, our EIF are Neyman-orthogonal, first-order nuisance errors cancel and we have second order bias remainder

$$R_n(\beta_B) = E[\hat{\mathbb{IF}}(\beta_B)] - E[\mathbb{IF}(\beta_B)].$$

By definition of the EIF, this can be decomposed as

$$R_n(\beta_B) = R_n(N_0^+) + R_n(N_0^-) + R_n(N_{B,1}^+) + R_n(N_{B,1}^-) - \beta_B\, R_n(\pi_{ac}),$$

We now bound all these components. Let $O = (SY, Y, S, D, Z, X)$ denote the data. Using the explicit form of the EIF, note that each term of $D^2\Psi[\eta_t](\Delta\eta, \Delta\eta)$ can be written as a finite sum of expressions of the following form:

$$\begin{aligned}
&\text{(A)} \quad E\big[U_t(O)\,\Delta a(O)\,\Delta b(O)\big],\\
&\text{(B)} \quad E\big[U_t(O)\,\Delta a(O)\,\{G(Q(X)+\Delta Q(X)) - G(Q(X))\}\big],\\
&\text{(C)} \quad E\big[U_t(O)\,\Delta\varepsilon(O)\big],
\end{aligned}$$

where

- $U_t(O) \in L^2$ are square integrable or bounded, components of the EIF (e.g., $W(Z,X)[DS - r(Z,X)]$, $W(Z,X)[(1-D)S - m(Z,X)]$ and similar).
- $\Delta a, \Delta b$ are primitive nuisance errors (one of $\Delta\mu, \Delta\nu, \Delta m, \Delta r, \Delta e, \Delta F, \Delta G_L, \Delta G_U$).
- $G \in \{G_L, G_U\}$ are the trimmed expectations.
- Q_1 is a true trimming or quantile boundary, and ΔQ_1 is the associated boundary error.
- $\Delta\varepsilon(O)$ are of the classification error type arising from the lack of knowledge of the positive or negative monotonicity direction, i.e., proportional to $\mathbb{1}(\lambda_0(x) < \lambda_1(x)) - \mathbb{1}(\hat\lambda_0(x) < \hat\lambda_1(x))$ or with reversed sign.

By Cauchy–Schwarz and the local Lipschitz property of G near the boundary guaranteed by Assumption A.3 and A.4, we will show that

$$|E[U_t\,\Delta a\,\Delta b]| \le \|U_t\|_2\,\|\Delta a\|_2\,\|\Delta b\|_\infty,$$

and

$$|E[U_t\,\Delta a\,\{G(Q_1+\Delta Q_1) - G(Q_1)\}]| \le \|U_t\|_2\,\|\Delta a\|_2\,L_G\,\|\Delta Q_1\|_\infty.$$

Moreover, in the following we present a quantile bound

$$\|\Delta Q_1\|_\infty \lesssim \|\Delta F\|_{\infty,\mathcal{N}} + \|\Delta m\|_\infty + \|\Delta r\|_\infty,$$

and analogously for the "minus" quantile. Therefore every $\|\Delta Q_1\|_\infty$ can be bounded by a combination of the primitive nuisance errors. Overall, we show that the total remainder is bounded via mixed L^2 and (localized) L^∞ norms of the primitive nuisance components, and classification error, i.e., of the shape

$$|R_n(\beta_B)| = |\Psi(\hat{\eta}) - \Psi(\eta)| \lesssim \sum_{(i,j)} C_{ij} \, \|\Delta\eta_i\|_{p_i} \, \|\Delta\eta_j\|_{p_j} \; + \; R_{\text{cls}} \; + \; o_p(n^{-1/2}),$$

where C_{ij} are finite constants only depending on the joint distribution of observables and $R_{\text{cls}} \lesssim (\|\Delta m\|_\infty + \|\Delta r\|_\infty)^{\kappa+1}$ is a margin-dependent classification remainder.

C.2.2. Remainder Decomposition

By the stabilized score and orthogonality,

$$R_n(\beta_B) = R_n(N_0^+) + R_n(N_0^-) + R_n(N_{B,1}^+) + R_n(N_{B,1}^-) - \beta_B \, R_n(\pi_{ac}),$$

where each block remainder is the quadratic (or classification) part arising from plugging $\hat{\eta}$ into their respective influence function $\mathbb{IF}(\cdot)$. Concretely:

- $R_n(N_0^\pm)$: products of errors for μ, ν and e plus classification errors from using estimated $\mathbb{1}^\pm$.
- $R_n(N_{B,1}^+)$: uses $\psi_B(z,x)$, Q_1, F_1, m, r, e plus classification errors from using estimated $\mathbb{1}^+$.
- $R_n(N_{B,1}^-)$: uses $\psi_B^-(z,x)$, Q_0, F_0, m, r, e plus classification errors from using estimated $\mathbb{1}^-$.
- $R_n(\pi_{ac})$: quadratic in errors for m, r plus classification errors from using estimated $\mathbb{1}^\pm$.

We now provide detailed bounds in terms of the nuisance estimation errors. For any derived nuisances, we reduce them to their primitives as given in Table 4.1.

C.2.3. Bahadur Expansion and Quantile Deviation Bound

Key Identity for the True CDF

By Assumption A.4, for any $y \in \mathcal{N}_x$,

$$\big|F_1(y|x) - F_1(Q_1(x)|x)\big| \;=\; \Big|\int_{Q_1(x)}^{y} f_1(t|x)\,dt\Big| \;\geq\; f_{\min}\,|y - Q_1(x)|. \tag{C.1}$$

Triangle Inequality and Monotonicity

By definition of quantiles, $F_1(Q_1(x)|x) = p(x)$ and $\hat{F}_1(\hat{Q}_1(x)|x) = \hat{p}(x)$. Hence

$$F_1(\hat{Q}_1(x)|x) - F_1(Q_1(x)|x) = \left[F_1(\hat{Q}_1(x)|x) - \hat{F}_1(\hat{Q}_1(x)|x)\right] + \left[\hat{p}(x) - p(x)\right]$$

Taking absolute values and using the local sup-norm on $\mathcal{N}_x$,

$$\left|F_1(\hat{Q}_1(x)|x) - F_1(Q_1(x)|x)\right| \;\le\; \sup_{y\in\mathcal{N}_x} \left|\hat{F}_1(y|x) - F_1(y|x)\right| \;+\; |\hat{p}(x) - p(x)|. \tag{C.2}$$

Inversion

Applying (C.1) with $y = \hat{Q}_1(x)$ and combining with (C.2) yields

$$f_{\min} \, |\hat{Q}_1(x) - Q_1(x)| \;\le\; \sup_{y\in\mathcal{N}_x} \left|\hat{F}_1(y|x) - F_1(y|x)\right| \;+\; |\hat{p}(x) - p(x)|.$$

Thus,

$$|\hat{Q}_1(x) - Q_1(x)| \;\le\; \frac{1}{f_{\min}} \left(\sup_{y\in\mathcal{N}_x} \left|\hat{F}_1(y|x) - F_1(y|x)\right| \;+\; |\hat{p}(x) - p(x)| \right). \tag{C.3}$$

Neighborhood Localization

Now if

$$\sup_{y\in\mathcal{N}_x} \left|\hat{F}_1(y|x) - F_1(y|x)\right| + |\hat{p}(x) - p(x)| \xrightarrow{p} 0$$

uniformly in x, then for large n bound (C.3) implies $|\hat{Q}_1(x) - Q_1(x)| \le \xi$ with probability $\to 1$, so $P(\hat{Q}_1(x) \in \mathcal{N}_x) \to 1$ and (C.3) is valid.

Uniform Bound over x

Recall the uniform local norm $\|\Delta F\|_{\infty,\mathcal{N}} := \sup_x \sup_{y\in\mathcal{N}_x} |\hat{F}_1(y|x) - F_1(y|x)|$. If

$$|\hat{p}(x) - p(x)| \;\lesssim\; \left(\|\Delta m\|_\infty + \|\Delta r\|_\infty\right),$$

then taking $\sup_x$ in (C.3) yields

$$\|\Delta Q_1\|_\infty \;\lesssim\; \frac{1}{f_{\min}} \left(\|\Delta F\|_{\infty,\mathcal{N}} + \left(\|\Delta m\|_\infty + \|\Delta r\|_\infty\right)\right). \tag{C.4}$$

Note that no differentiability of $\hat{F}_1$ is required but only monotonicity of $\hat{F}_1$ and the structural lower bound $f_{\min}$ for F_1 near $Q_1(x)$. Applying the same argument to the negative equivalents with F_0, Q_0, yields

$$\|\Delta Q_0\|_\infty \lesssim \frac{1}{f_{\min}} \left(\|\Delta F\|_{\infty,\mathcal{N}} + \|\Delta m\|_\infty + \|\Delta r\|_\infty\right). \tag{C.5}$$

as the norms are also uniform over the conditional arguments z, d.

C.2.4. Trimmed Expectations Bound

We now show the bound for the trimmed expectations for the positive monotonicity part of the lower bound case and then for the negative part. The upper bound follows analogously. We suppress the irrelevant components to ease notation of the trimmed expectation $\psi_L(z,x) = G_L(Q_1(p(X),X)|1,z,x) = G_L(Q_1(x)|x)$ and its estimator $\hat{\psi}_L(z,x) = \hat{G}_L(\hat{Q}_1(p(\hat{X}),X)|1,z,x) = \hat{G}_L(\hat{Q}(x)|x)$. Decompose

$$\begin{aligned}\Delta\psi_L(z,x) &= \hat{G}_L(\hat{Q}_1(x)|x) - G_L(Q_1(x)|x)\\ &= \underbrace{\left[\hat{G}_L(Q_1(x)|x) - G_L(Q_1(x)|x)\right]}_{(I)} + \underbrace{\left[\hat{G}_L(\hat{Q}_1(x)|x) - \hat{G}_L(Q_1(x)|x)\right]}_{(II)}.\end{aligned} \tag{C.6}$$

Term (I) is a pointwise error which can be bounded by the local worst-case

$$|(I)| \le \sup_z \sup_x \sup_{|y-Q_1(x)|\le\xi} |\hat{G}_L(y|x) - G_L(y|x)| = \|\Delta G\|_{\infty,\mathcal{N}}.$$

For (II), adding and subtracting $\hat{Q}_1$ at the true threshold and applying the triangle inequality plus Assumption A.4 yields:

$$\begin{aligned}|(II)| &\le |\hat{G}_L(\hat{Q}_1(x)|x) - G_L(\hat{Q}_1(x)|x)| + |G_L(\hat{Q}_1(x)|x) - G_L(Q_1(x)|x)|\\ &\le \|\Delta G\|_{\infty,\mathcal{N}} + L_G\,|\hat{Q}_1(x) - Q_1(x)|.\end{aligned} \tag{C.7}$$

Combining (C.6)–(C.7) and taking the supremum over (z,x) then yields

$$\|\Delta\psi_L\|_\infty \le 2\,\|\Delta G\|_{\infty,\mathcal{N}} + L_G\,\|\Delta Q_1\|_\infty. \tag{C.8}$$

By the same argument with G_U in place of G_L, we also obtain

$$\|\Delta\psi_U\|_\infty \le 2\,\|\Delta G_U\|_{\infty,\mathcal{N}} + L_G\,\|\Delta Q_1\|_\infty. \tag{C.9}$$

For the negative monotonicity side, we equivalently can derive

$$\|\Delta\psi_L^-\|_\infty \le 2\,\|\Delta G_L\|_{\infty,\mathcal{N}} + L_G\,\|\Delta Q_0\|_\infty, \tag{C.10}$$

$$\|\Delta\psi_U^-\|_\infty \le 2\,\|\Delta G_U\|_{\infty,\mathcal{N}} + L_G\,\|\Delta Q_0\|_\infty, \tag{C.11}$$

as all the norms are also supremum over d. Thus, using the quantile deviation bounds (C.4)–(C.5) inside (C.8)–(C.11) yields full control of $\Delta\psi$ solely in terms of primitives $\Delta F, \Delta G$, and $\Delta m, \Delta r$ (through Δp), together with structural constants $f_{\min}$ and L_G.

C.2.5. Margin-partition Remainder for π_{ac}

We now control the second type of component that arises from boundary displacement/classification error in the density. The identical term arises in the classification error in $\mathbb{1}^\pm$ that enters N^0 and $N^\pm_{B,1}$. Define the contrast of the arguments and its estimator as

$$\Gamma(x) := \lambda_0(x) - \lambda_1(x), \qquad \hat{\Gamma}(x) := \hat{\lambda}_0(x) - \hat{\lambda}_1(x).$$

Let the uniform estimation error be

$$\varepsilon := \|\Delta\Gamma\|_\infty = \sup_x |\hat{\Gamma}(x) - \Gamma(x)|.$$

The true and estimated positive regions are then given by

$$\mathcal{X}^+ := \{x : \Gamma(x) \le 0\}, \qquad \hat{\mathcal{X}}^+ := \{x : \hat{\Gamma}(x) \le 0\}.$$

Also denote their complements as $\mathcal{X}^{+c}$ and $\hat{\mathcal{X}}^{+c}$ respectively. The symmetric difference between the sets is then

$$\hat{\mathcal{X}}^+ \oplus \mathcal{X}^+ = \left(\mathcal{X}^+ \cap \hat{\mathcal{X}}^{+c}\right) \cup \left(\mathcal{X}^{+c} \cap \hat{\mathcal{X}}^+\right).$$

We now show that

$$\hat{\mathcal{X}}^+ \oplus \mathcal{X}^+ \subseteq \{x : |\Gamma(x)| \le \varepsilon\}.$$

Case 1: Let $x \in \mathcal{X}^+ \cap \hat{\mathcal{X}}^{+c}$. Then $\Gamma(x) \le 0$ and $\hat{\Gamma}(x) > 0$. Hence

$$0 < \hat{\Gamma}(x) = \Gamma(x) + (\hat{\Gamma}(x) - \Gamma(x)) \Rightarrow -\Gamma(x) < \hat{\Gamma}(x) - \Gamma(x).$$

Since $\Gamma(x) \le 0$, we have $|\Gamma(x)| = -\Gamma(x)$, and therefore

$$|\Gamma(x)| \le |\hat{\Gamma}(x) - \Gamma(x)| \le \|\Delta\Gamma\|_\infty = \varepsilon.$$

Case 2: Let $x \in \mathcal{X}^{+c} \cap \hat{\mathcal{X}}^{+}$. Then $\Gamma(x) > 0$ and $\hat{\Gamma}(x) \leq 0$. Hence

$$0 \geq \hat{\Gamma}(x) = \Gamma(x) + (\hat{\Gamma}(x) - \Gamma(x)) \Rightarrow \Gamma(x) \leq \Gamma(x) - \hat{\Gamma}(x) = |\hat{\Gamma}(x) - \Gamma(x)|.$$

Since $\Gamma(x) > 0$, we have $|\Gamma(x)| = \Gamma(x)$, and therefore

$$|\Gamma(x)| \leq |\hat{\Gamma}(x) - \Gamma(x)| \leq \|\Delta\Gamma\|_\infty = \varepsilon.$$

Therefore we obtain that the symmetric difference set is subset of a small epsilon set for the population difference

$$\widehat{\mathcal{X}}^{+} \oplus \mathcal{X}^{+} \;\subseteq\; \{x : \; |\lambda_0(x) - \lambda_1(x)| \leq \epsilon\}$$

and thus, by margin assumption A.5

$$P\big(\widehat{\mathcal{X}}^{+} \oplus \mathcal{X}^{+}\big) \leq C_M\, \epsilon^{\kappa}.$$

Overall, since each misclassification multiplies an L^∞ integrand error $O(\epsilon)$, this yields overall classification remainder

$$R_{\text{cls}} \;\lesssim \big(\|\Delta m\|_\infty + \|\Delta r\|_\infty\big)^{\kappa+1}. \tag{C.12}$$

C.2.6. Bounding EIF Block Remainders and $R_n(\beta_B)$

We now obtain the remainders for all objects of type (A), (B) and (C). $R_n(N_0^+), R_n(N_0^-)$ contain no quantiles and are of classical "augmented IPW" (AIPW) form interacted with different mean functions μ and ν respectively and a classification indicator. Thus, their second order remainder is bounded by

$$|R_n(N_0^+)| + |R_n(N_0^-)| \;\lesssim\; \|\Delta\mu\|_2\, \|\Delta e\|_2 \;+\; \|\Delta\nu\|_2\, \|\Delta e\|_2 \;+\; R_{\text{cls}}.$$

For the positive side mixture component error $R_n(N_{B,1}^+)$ we have a product form with nuisances $\psi_B = G_B(Q_1), Q_1, F, m, r, e$ and classification indicator. Thus, we obtain that

$$\begin{aligned}
&|R_n(N_{B,1}^+)| \\
&\lesssim (||\Delta e||_2 + ||\Delta m||_2 + ||\Delta r||_2) \times \Big[||\Delta e||_2 + ||\Delta m||_2 + ||\Delta r||_2 + ||\Delta G||_{\infty,\mathcal{N}} + ||\Delta F||_{\infty,\mathcal{N}} + ||\Delta Q_1||_\infty\Big] \\
&\lesssim (||\Delta e||_2 + ||\Delta m||_2 + ||\Delta r||_2) \times \Big[||\Delta e||_\infty + ||\Delta m||_\infty + ||\Delta r||_\infty + ||\Delta G||_{\infty,\mathcal{N}} + ||\Delta F||_{\infty,\mathcal{N}}\Big],
\end{aligned}$$

as $||a||_2 \lesssim ||a||_\infty$ as well as the upper bound for $||\Delta Q_1||_\infty$ from (C.4). Now note that the same bound applies to $|R_n(N_{B,1}^-)|$. Moreover, for the always-selected complier density we have the mixture probability nuisances as well as classification in the remainder:

$$|R_n(\pi_{ac})| \;\lesssim\; \left(\|\Delta m\|_2^2 + \|\Delta r\|_2^2\right) \;+\; R_{\text{cls}}.$$

The rate for R_{cls} is given in (C.12). Combining this with the primitive rates for the remaining elements as well as the result from Section B.4 and the fact that β_B is finite, then yields total bound the remainder

$$\begin{aligned}|R_n(\beta_B)| \lesssim{}& \|\Delta\mu\|_2\,\|\Delta e\|_2 \;+ \|\Delta\nu\|_2\,\|\Delta e\|_2 \;+ (||\Delta m||_\infty + ||\Delta r||_\infty)^{\kappa+1} \\ &+ (||\Delta e||_2 + ||\Delta m||_2 + ||\Delta r||_2) \times \Big[||\Delta e||_\infty + ||\Delta m||_\infty + ||\Delta r||_\infty + ||\Delta G||_{\infty,\mathcal{N}} + ||\Delta F||_{\infty,\mathcal{N}}\Big],\end{aligned} \tag{C.13}$$

with B arbitrary.

C.3. Proof of Theorem 4.1

We now denote $E_n[X] = \frac{1}{n}\sum_i^n X_i$ and $G_n[X] = \frac{1}{\sqrt{n}}\sum_i^n (X_i - E[X_i])$. Recall that the bounds are estimated via

$$\hat{\beta}_B = \frac{\hat{N}_0^+ + \hat{N}_0^- + \hat{N}_{B,1}^+ + \hat{N}_{B,1}^-}{\hat{\pi}_{ac}},$$

where all estimators are obtained by solving their respective influence functions at the empirical solution with cross-fitted nuisances. Thus $\hat{N}_0$ is defined by solving $E_n[\mathbb{IF}(\hat{N}_0, \hat{\eta})] = 0$ and equivalently for the remaining parameters. Thus, using standard Slutzky arguments, we obtain the following linearization

$$\sqrt{n}(\hat{\beta}_B - \beta_B) = G_n[\hat{\mathbb{IF}}(\beta_B)]\bigg(1 + O\bigg(\sup_{N\in\{N_0^+, N_0^-, N_1^+, N_1^-\}} |E_n[\hat{\mathbb{IF}}(N)] - E[\mathbb{IF}(N)]|\bigg)\bigg).$$

Further decomposing the leading term yields

$$\begin{aligned}G_n[\hat{\mathbb{IF}}(\beta_B)] ={}& G_n[\mathbb{IF}(\beta_B)] + (G_n[\hat{\mathbb{IF}}(\beta_B)] - G_n[\mathbb{IF}(\beta_B)]) \\ &+ \sqrt{n}E[\hat{\mathbb{IF}}(\beta_B) - \mathbb{IF}(\beta_B)].\end{aligned}$$

As nuisances are cross-fitted, the empirical process is $o_p(1)$ as long as the nuisances are consistent as implied by Assumption A.7. The big $O(\cdot)$ remainder will be $O_p(n^{-1/2})$

as their $\sqrt{n}$-analogues are already $O_p(1)$ as implied by the results in Appendix B. Thus, overall, we obtain that

$$\sqrt{n}(\hat{\beta}_B - \beta_B) = G_n[\mathbb{IF}(\beta_B)] + o_p(1) \xrightarrow{d} \mathcal{N}(0, E[\mathbb{IF}(\beta_B)^2]).$$

Here $E[\mathbb{IF}(\beta_B)^2]$ is the semiparametric efficiency bound for β_B as we have explicitly used the Riesz representer of the pathwise derivative under the nonparametric model and regularity conditions, see Bickel et al. (1993) or Kennedy (2024). Moreover, since the leading expression of each bound is asymptotically linear, we also obtain joint convergence of $\sqrt{n}(\hat{\beta} - \beta)$ by Assumption A.2 and the Cramer–Wold device.

C.4. Proof of Proposition 4.1

We now show the equivalence to existing results under additional restrictions. We focus on the lower bound under strong sample selection monotonicity, but the analogous derivations apply to the weak sample selection monotonicity case and the upper bound as well via simple sign/partition adjustment as in Section C.1.3. Statements about equality of random variables are almost surely. Note that, in this case we denote $N_B^+ = N_B$ and have $N_B^- = 0$.

C.4.1. (i) $Z = D$: Lee Bounds

Denote $s(d, x) = P(S = 1|D = d, X = x)$ and

$$\beta_{1,d}(x, u) = E[Y|S = 1, D = d, X = x, Y \leq Q_1(u, x)],$$
$$\beta_{0,d}(x, u) = E[Y|S = 1, D = d, X = x, Y \geq Q_1(u, x)],$$

which yields

$$\beta_{1,d}(x, 1) = \beta_{0,d}(x, 0) = E[Y|S = 1, D = d, X = x].$$

When $Z = D$ (perfect compliance), the nuisances simplify as follows

$$P(Z=1|X) = P(D=1|X) = e(x),$$

$$r(z,x) = E[DS|Z=z, X=x] = \begin{cases} 0 & \text{if } Z=D=0 \\ s(1,x) & \text{if } Z=D=1 \end{cases},$$

$$m(z,x) = E[S|Z=z, X=x] - E[DS|Z=z, X=x] = \begin{cases} s(0,x) & \text{if } Z=D=0 \\ 0 & \text{if } Z=D=1 \end{cases},$$

$$\mu(z,x) = E[Y(1-D)S|Z=z, X=x] = \begin{cases} 0 & \text{if } Z=D=1 \\ E[Y|S=1, D=0, X=x]s(0,x) & \text{if } Z=D=0 \end{cases},$$

$$\begin{aligned} F(y|d,z,x) &= P(Y \leq y|D=d, S=1, Z=z, X=x) \\ &= \begin{cases} 0 & \text{if } Z \neq D \\ P(Y \leq y|DS=1, X=x) & \text{if } Z=D \end{cases}. \end{aligned}$$

These yield the following simplifications

$$\begin{aligned} F_{Y_1|S_1=1, D_1>D_0, X}(y|x) &= \frac{F(y|1,1,x)r(1,x) - F(y|1,0,x)r(0,x)}{r(1,x) - r(0,x)} \\ &= F(y|1,1,x) \\ &= P(Y \leq y|DS=1, Z=1, x) \\ &=: F(y|1,x), \\ p(x) &= -\frac{(m(1,X) - m(0,X))}{(r(1,X) - r(0,X))} \\ &= \frac{s(0,x)}{s(1,x)}, \\ W(z,x) = W(d,x) &= \frac{d}{P(D=1|X=x)} - \frac{1-d}{1-P(D=1|X=x)}. \end{aligned}$$

Lastly, for the trimmed mean we have

$$\begin{aligned} &\psi_L(z,x) = E[Y\mathbb{1}(Y \leq Q_1(p(X),X))|DS=1, Z=z, X=x] \\ &= \begin{cases} 0 & \text{if } Z=D=0 \\ E[Y|Y \leq Q_1(p(X),X), DS=1, X=x]P(Y \leq Q_1(p(X),X)|DS=1, X=x) & \text{if } Z=D=1 \end{cases} \\ &= \begin{cases} 0 & \text{if } Z=D=0 \\ \beta_{1,1}(x, p(x))p(x) & \text{if } Z=D=1 \end{cases}. \end{aligned}$$

We now look at the three components of the influence functions for β_L. First, we start with π_{ac}. Noting that $Z=D$ we have that

$$
\begin{aligned}
\mathbb{IF}(\pi_{ac}) + \pi_{ac} &= -\left[W(Z,X)((1-D)S - m(Z,X)) + m(1,X) - m(0,X)\right] \\
&= -\left[\left(\frac{D}{e(X)} - \frac{1-D}{1-e(X)}\right)((1-D)S - m(D,X)) + m(1,X) - m(0,X)\right] \\
&= \frac{(1-D)}{(1-e(X))}(S - s(0,x)) + s(0,x),
\end{aligned}
$$

as $m(1,x) = 0$ and thus $Dm(D,X) = 0$. This is exactly the denominator in Heiler et al. (2024), Table 6.1. We now consider the numerator given by $N_1 + N_0$.

Now for the N_0 component we exploit that $\mu(1,x) = 0$ and thus $D\mu(D,X) = 0$. This yields

$$
\begin{aligned}
\mathbb{IF}(N_0) + N_0 &= W(Z,X)((1-D)SY - \mu(Z,X)) + \mu(1,X) - \mu(0,X) \\
&= \left(\frac{D}{e(X)} - \frac{1-D}{1-e(X)}\right)((1-D)SY - \mu(D,X)) + \mu(1,X) - \mu(0,X) \\
&= -\left[\frac{(1-D)}{(1-e(X))}SY - \mu(0,X)\left(1 - \frac{(1-D)}{(1-e(X))}\right)\right] \\
&= -\left[\frac{(1-D)}{(1-e(X))}SY - s(0,x)\beta_{0,0}(x,0)\left(1 - \frac{(1-D)}{(1-e(X))}\right)\right].
\end{aligned}
$$

Now for N_1 we first denote the following auxiliary results:

(N1.i)

$$
W(Z,X)(DS - r(Z,X))\psi_L(Z,X) = \frac{D}{e(X)}(S - s(1,X))\beta_{1,1}(X,p(X))p(X).
$$

(N1.ii)

$$
\begin{aligned}
&W(Z,X)DS(Y\mathbb{1}(Y \leq Q_1(p(X),X)) - \psi_L(Z,X)) \\
&\quad = \frac{DS}{e(X)}Y\mathbb{1}(Y \leq Q_1(p(X),X)) - \frac{DS}{e(X)}\beta_{1,1}(X,p(X))p(X).
\end{aligned}
$$

(N1.iii)

$$
\begin{aligned}
\psi_L(1,X)r(1,X) - \psi_L(0,X)r(0,X) &= \psi_L(1,X)r(1,X) \\
&= \beta_1(X,p(X))p(X)s(1,X) \\
&= \beta_1(X,p(X))s(0,X).
\end{aligned}
$$

(N1.iv) Note that as when $D = Z = 1$ we have $F_1(y|1,x) = P(Y \leq y|DS = 1, Z = 1, x) =: F(y|1,x)$ and same for the corresponding quantiles. Thus

$$-Q_1(p(X),X)W(Z,X)\Big[((1-D)S-m(z,X))$$
$$+(DS-r(Z,X))F(Q_1(p(X),X)|Z,X)+DS(\mathbb{1}(Y\leq Q_1(p(X),X))-F(Q_1(p(X),X)|Z,X))\Big]$$
$$=-Q_1(p(X),X)\Big(-\left[\frac{(1-D)}{1-e(X)}(S-s(0,X))\right]+\frac{D}{e(X)}(S-s(1,X))F_1(Q_1(p(X),X)|1,X)$$
$$+\frac{DS}{e(X)}(\mathbb{1}(Y\leq Q_1(p(X),X)-F_1(Q_1(p(X),X)|1,X))\Big)$$
$$=-Q_1(p(X),X)\left(\frac{DS}{e(X)}(\mathbb{1}(Y\leq Q_1(p(X),X)-p(X))\right)$$
$$+Q_1(p(X),X)\left(\frac{(1-D}{1-e(X)}(S-s(0,X))-\frac{D}{e(X)}(S-s(1,X))p(X)\right).$$

Summing up $(N1.i)$ to $(N1.iv)$ and adding the results of N_0 then yields

$$\begin{aligned}\mathbb{IF}(N_1)+N_1+\mathbb{IF}(N_0)+N_0=&\frac{DS}{e(X)}Y1\{Y\leq Q_1(p(X),X)\}-\frac{(1-D)S}{1-e(X)}Y\\&-\frac{DS}{e(X)}Q_1(p(X),X)\left[1\{Y\leq Q_1(p(X),X)\}-p(X)\right]\\&+Q_1(p(X),X)\left[\frac{1-D}{1-e(X)}(S-s(0,X))-p(X)\frac{D}{e(X)}(S-s(1,X))\right]\\&+s(0,X)\left[\beta_{1,1}(X,p(X))\left(1-\frac{D}{e(X)}\right)-\beta_{0,0}(X,0)\left(1-\frac{1-D}{1-e(X)}\right)\right]\end{aligned}$$

,

which is exactly the corresponding component of the influence function proposed in Heiler et al. (2024), Table 6.1., see also Semenova (2025), Section 6.6.

C.4.2. (ii) $S=1$: LATE

If $S=1$, we have the following nuisance simplifications:

$$r(z,x)=E[DS|Z=z,X=x]=E[D|Z=z,X=x],$$
$$m(z,x)=E[(1-D)S|Z=z,X=x]=E[(1-D)|Z=z,X=x].$$

Thus we have that $r(z,x)=1-m(z,x)$ and thus $-(m(1,x)-m(0,x))=r(1,x)-r(0,x)$ which implies that $p(x)=1$. Moreover, we have that $F(y|d,z,x)=P(Y\leq y|D=1,Z=z,X=x)$ (without S).

Auxiliary Result: CDFs at $Q_1(1,X)$ We now show that when the trimming threshold $p=1$, the conditional cdfs must also be equal to one. We omit x for simplicity. In particular, $F(y|z)$ is a continuous CDF with $\lim_{y\to-\infty}F(y|z)=0$ and $\lim_{y\to+\infty}F(y|z)=1$. Define generic

$$b_z \;=\; \inf\{\, y : F(y|z) = 1\}, \qquad F_1(y) \;=\; \frac{r_1\, F(y|1) \;-\; r_0\, F(y|0)}{r_1 - r_0}, \qquad Q_1 \;=\; \inf\{\, y : F_1(y) = 1\}.$$

Then for $y < \max(b_0, b_1)$, one has $F_1(y) < 1$ and hence $Q_1 \geq \max(b_0, b_1)$. Conversely, for $y \geq \max(b_0, b_1)$, $F(y|0) = F(y|1) = 1$ implies $F_1(y) = 1$ and thus $Q_1 \leq \max(b_0, b_1)$. Therefore$Q_1 = \max(b_0, b_1)$. Lastly, since $Q_1 \geq b_z$, continuity gives $F(Q_1|z) = 1 \quad \text{for } z = 0, 1$. In the original notation this means that almost surely

$$F(Q_1(p(X), X)|Z, X) = 1.$$

We now move the components of the influence function. First regarding π_{ac} we obtain the following simplification

$$\begin{aligned}
\mathbb{IF}(\pi_{ac}) + \pi_{ac} &= -\left[W(Z,X)((1-D) - m(Z,X)) + m(1,X) - m(0,X)\right] \\
&= W(Z,X)(D - r(Z,X)) + r(1,X) - r(0,X) \\
&= \frac{Z}{P(Z=1|X)}(D - r(1,X)) - \frac{(1-Z)}{1-P(Z=1|X)}(D - r(0,X)) + r(1,X) - r(0,X),
\end{aligned}$$

which is exactly the classic uncentered AIPW EIF of the share of compliers. Now for N_0, we obtain

$$\mathbb{IF}(N_0) + N_0 = W(Z,X)((1-D)Y - \mu(Z,X)) + \mu(1,X) - \mu(0,X),$$

while for N_1 we first note that

$$\begin{aligned}
((1-D)S - m(Z,X)) + (DS - r(Z,X))F(Q_1(p(X),X)|Z,X) \\
+ DS(\mathbb{1}(Y \leq Q_1(p(X),X)) - F(Q_1(p(X),X)|Z,X)) \\
&= ((1-D) - m(Z,X)) + (D - r(Z,X) + 0 \\
&= 1 - m(Z,X) - r(Z,X) \\
&= 0.
\end{aligned}$$

Thus we have that

$$\begin{aligned}
\mathbb{IF}(N_1) + N_1 &= W(Z,X)(D - r(Z,X))\psi_L(Z,X) + W(Z,X)D[Y - \psi_L(Z,X)] \\
&\quad + \psi_L(1,X)r(1,X) - \psi_L(0,X)r(0,X) \\
&= W(Z,X)[DY - \psi_L(Z,X)r(Z,X)] + \psi_L(1,X)r(1,X) - \psi_L(0,X)r(0,X).
\end{aligned}$$

Now note that, by definition

$$
\begin{aligned}
\psi(z,x)r(z,x)+\mu(z,x) &= E[Y|D=1,Z=z,X=x]E[D|Z=z,X]+E[Y(1-D)|Z=z,X=x] \\
&= E[YD|Z=z,X=x]+E[Y(1-D)|Z=z,X=x] \\
&= E[Y|Z=z,X=x].
\end{aligned}
$$

Thus, for the total numerator we obtain

$$
\begin{aligned}
&\mathbb{IF}(N_1)+N_1+\mathbb{IF}(N_0)+N_0 \\
&\quad= W(Z,X)[DY-\psi_L(Z,X)r(Z,X)]+\psi_L(1,X)r(1,X)-\psi_L(0,X)r(0,X) \\
&\qquad+W(Z,X)[(1-D)Y-\mu(Z,X)]+\mu(1,X)-\mu(0,X) \\
&\quad= W(Z,X)Y+[\psi_L(1,X)r(1,X)+\mu(1,X)]\left(1-\frac{Z}{P(Z=1|X)}\right) \\
&\qquad-[\psi_L(0,X)r(0,X)+\mu(0,X)]\left(1-\frac{(1-Z)}{1-P(Z=1|X)}\right) \\
&\quad= W(Z,X)Y+E[Y|Z=1,X]\left(1-\frac{Z}{P(Z=1|X)}\right)-E[Y|Z=0,X]\left(1-\frac{(1-Z)}{1-P(Z=1|X)}\right) \\
&\quad= \frac{Z}{P(Z=1|X)}(Y-E[Y|Z=1,X])-\frac{(1-Z)}{1-P(Z=1|X)}(Y-E[Y|Z=0,X]) \\
&\qquad+E[Y|Z=1,X]-E[Y|Z=0,X],
\end{aligned}
$$

which is exactly the EIF for the numerator/reduced form component of the LATE. Combining results with the denominator yield the well-known EIF for the LATE (Frölich, 2007; Chernozhukov et al., 2018; Heiler, 2022).

C.4.3. (iii) $Z=D, S=1$: ATE

In this case we have all the $S=1$ simplifications as well as $Z=D$ and thus $ZD=D$, $(1-Z)D=0$, and

$$
\begin{aligned}
r(0,x) &= m(1,x) = 0, \\
r(1,x) &= m(0,x) = 1.
\end{aligned}
$$

This yields

$$
\begin{aligned}
\mathbb{IF}(\pi_{ac}) &= \frac{D}{P(D=1|X)}(D-r(1,X))-\frac{(1-D)}{1-P(D=1|X)}(D-r(0,X))+r(1,X)-r(0,X) \\
&= 1
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{IF}(N_1) + N_1 + \mathbb{IF}(N_0) + N_0 \\
&\quad = \frac{D}{P(D=1|X)}(Y - E[Y|D=1,X]) - \frac{(1-D)}{1-P(D=1|X)}(Y - E[Y|D=0,X]) \\
&\qquad + E[Y|D=1,X] - E[Y|D=0,X],
\end{aligned}
$$

which is the usual efficient AIPW influence function for the ATE (Hahn, 1998).

D. Monte Carlo Simulations

D.1. Design

In this section we report a simple Monte Carlo design to assess coverage and power properties of the suggested confidence intervals in Section 4.3 in finite samples. The true DGP and its parameterization are in Equation (D.1) and Table D.1. It is a single index treatment response model with normal errors and an additional single index sample selection step similar to the designs in Heckman and Vytlacil (2005), Heiler (2022) and Bartalotti et al. (2023):

$$
\begin{aligned}
Z &\sim \text{Binom}(p_z) \\
X_1, \ldots, X_J &\sim \text{Unif}(0,1) \\
Y_1 &= \mu_y(1,X) + \varepsilon_1 \\
Y_0 &= \varepsilon_0 \\
D_z &= \mathbb{1}(U_d \le \mu_d(z,x)) \\
S_d &= \mathbb{1}(U_s \le \mu_s(d,x))
\end{aligned}
\qquad
\begin{pmatrix} \varepsilon_1 \\ \varepsilon_0 \\ U_d \\ U_s \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & 0 & 0 & \rho_{1s} \\ 0 & 1 & \rho_{0d} & 0 \\ 0 & \rho_{0d} & 1 & 0 \\ \rho_{1s} & 0 & 0 & 1 \end{pmatrix} \right) \tag{D.1}
$$

Table D.1: Monte Carlo: Design Parameters

Parameter/Function	Value
p_z	0.5
J	25
σ^2	1
ρ_{1s}	0.9999
ρ_{0d}	0.5
$\mu_y(1,x)$	$f_x/7.205 + \mathbb{1}(x_6 > 0.5) + \mathbb{1}(x_7 > 0.5)$
$\mu_d(z,x)$	$f_x/7.205 - 1 + 2z$
$\mu_s(d,x)$	$f_x - 0.5 + 2d$
f_x	$\sum_{j=1}^{5} x_j - 2.5$

Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and cumulative distribution function respectively. The design implies that the conditional SLATE is given

$$
\begin{aligned}
\theta_{SLATE}(x) &= E[Y_1 - Y_0|S_0 = 1, D_1 > D_0, X = x] \\
&= \mu_y(1,x) - \rho_{1,s}\frac{\phi(\mu_s(0,x))}{\Phi(\mu_s(0,x))} + \rho_{0,d}\frac{(\phi(\mu_d(1,x)) - \phi(\mu_d(0,x)))}{(\Phi(\mu_d(1,x)) - \Phi(\mu_d(0,x)))}.
\end{aligned}
$$

Moreover, the conditional share of always-selected compliers is

$$
\begin{aligned}
\pi_{ac}(x) &= P(S_0 = 1, D_1 > D_0|X = x) \\
&= \Phi(\mu_s(0,x))(\Phi(\mu_d(1,x)) - \Phi(\mu_d(0,x))).
\end{aligned}
$$

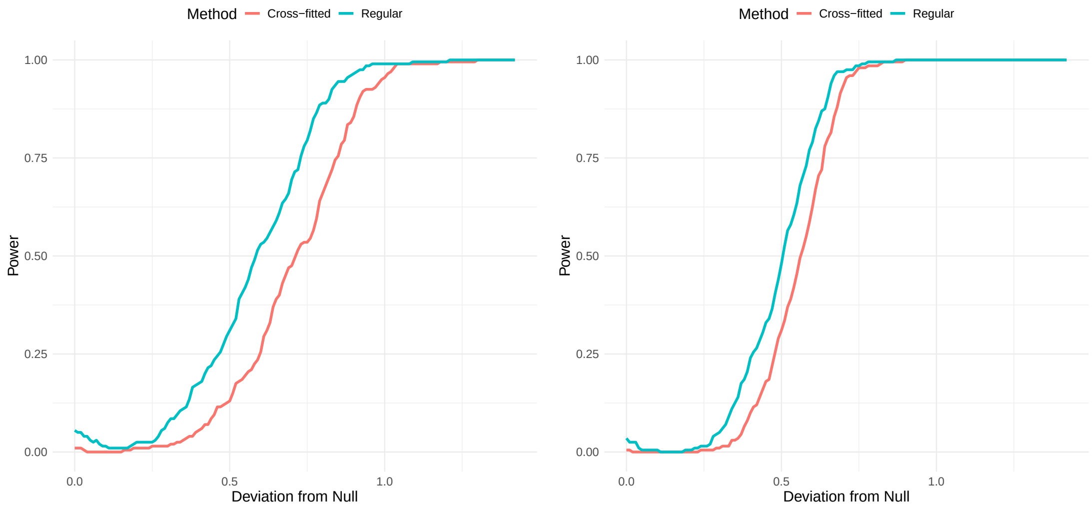
θ_{SLATE} can then be obtained as

$$
\theta_{SLATE} = \frac{\int \beta(x)\pi_{ac}(x)dP(x)}{\int \pi_{ac}(x)dP(x)}. \tag{D.2}
$$

For the Monte Carlo Simulations, our estimation follows the algorithm in Section 4.3. We estimate all nuisances with generalized random forests with default parameters, including the functional parameters which are estimated on a large grid. Strong sample selection monotonicity is not imposed, i.e., there is risk of missclassification of the response types. Before inversion of the conditional CDFs, we use isotonic regression for post-processing. The design is relatively sparse in the underlying nuisances such that convergence requirements as in Section 4.4 are credible. For evaluation, we report the average number of times where θ_{SLATE} is not contained in the confidence interval as a function of deviation from the true null with and without cross-fitting. At zero, this naturally corresponds to the size of the associated test.

Figure D.1 contains the power curves for both $n = 2000$ and $n = 5000$. Given the design, this corresponds to an effective sample size of approximately 600 and 1500 always-selected compliers respectively. Cross-fitting seems to be required to have the minimal rejection area close to the true null. The confidence intervals have conservative coverage at and close to the null. This is expected as the effect is not at the boundary of the identified set in line with, e.g., Stoye (2020) or Heiler (2024). For larger deviations power quickly increases. Overall, results suggest that the theory in Section 4.4 approximates the finite sample distribution well.

Figure D.1: Monte Carlo Study: Power Curves



Power curves as a function of deviation from the null for Sharp-DML bounds using cross-fitting according to Algorithm 1. Regular is equivalent but without cross-fitting. $n = 2000$ and $n = 5000$ observations and $M = 200$ replications.

E. Empirical Study II: Oregon Health Insurance Experiment

E.1. Data and Methods

In this section, we study the effects of Medicaid coverage on healthcare utilization in the Oregon Health Insurance Experiment (OHIE), following Finkelstein et al. (2012). In 2008, Oregon expanded Medicaid for low-income uninsured adults through a lottery because available funding was insufficient to cover all eligible applicants. Individuals selected by lottery obtained the opportunity to apply for coverage for themselves and household members, but actual enrollment required submitting paperwork and verifying eligibility, so take-up was incomplete.

We use the publicly available OHIE survey data analyzed by Finkelstein et al. (2012). The survey covers individuals from both lottery-selected and non-selected households and includes demographic information collected at sign-up. Our treatment is ever enrolling in Medicaid during the study period, while the instrument is the lottery-based assignment. We consider three outcomes: the number of outpatient visits, the number of prescription drugs, and log annual total medical expenditure. For outpatient visits and prescription drugs, outcomes are only observed for individuals with positive utilization, while for log expenditure the outcome is non-missing only when expenditure is positive. We additionally use pre-randomization covariates including gender, age, race and ethnicity, education, survey-wave fixed effects, and household-size fixed effects. The list of covariates along with sample summary statistics are provided in Table F.2.

The data are well suited to our framework. Lottery selection provides a credible source of exogenous assignment, while exclusion is plausible since effects on utilization should operate through actual Medicaid coverage rather than lottery selection itself. As in the JC application, noncompliance is substantial because only a minority of lottery winners ultimately enrolled, so assignment and treatment differ materially. Depending on the outcome, analysis samples range from 16,868 to 22,064 observations after restricting to units with non-missing treatment, instrument, outcome, and covariates.

We evaluate always-selected complier effects θ_{SLATE} for all three outcomes using (i) Chen and Flores (2015) bounds and (ii) sharp-basic bounds - both assuming strong sample selection monotonicity - as well as (iii) sharp DML bounds with covariates under weak sample selection monotonicity. Implementation of (i) follows Chen and Flores (2015). (ii) uses simple sample analogues of (2.19) and (2.20). (iii) is obtained via the procedure in Section 4.3, with the same implementation details as in the JC application.

E.2. Results: Bounds and Shares

Table E.1 reports results for outpatient visits, prescription drugs, and annual total expenditure. Across outcomes, the share of always-selected compliers is relatively small but non-negligible, ranging from about 15% to 21%. Sharp-DML is again the preferred specification because the data do not support strong sample selection monotonicity uniformly across outcomes. The estimated share of positive sample selection under Sharp-DML is 0.957 for outpatient visits, but only 0.626 for prescription drugs and 0.638 for expenditure, indicating that allowing the direction of sample selection to vary with covariates is empirically important, especially for the latter two outcomes.

First, we focus on the comparison between CF and Sharp-basic under strong sample selection monotonicity. The sharp bounds are substantially shorter than CF across all outcomes, with reductions in identified-set length of about 31.1% for outpatient visits, 50.4% for prescription drugs, and 69.4% for total expenditure. The corresponding

Table E.1: Bounds for the Effect of Health Insurance on Healthcare Utilization

	CF	Sharp-basic	Sharp-DML
Outpatient visits ($N = 22,064$)			
Bounds	$[-1.447,\ 1.520]$	$[-1.129,\ 0.915]$	$[-0.821,\ 1.597]$
Standard Errors	—	$(0.247,\ 0.333)$	$(0.299,\ 0.374)$
95% CI	$[-1.879,\ 2.157]$	$[-1.533,\ 1.460]$	$[-1.311,\ 2.211]$
Share of positive sample selection [a]	1 (assumed)		0.957 (0.004)
Share of *ac* [b]	0.147 (0.007)		0.145 (0.007)
Prescription drugs ($N = 17,223$)			
Bounds	$[-0.832,\ 0.125]$	$[-0.535,\ -0.060]$	$[-0.295,\ 0.113]$
Standard Errors	—	$(0.289,\ 0.244)$	$(0.257,\ 0.113)$
95% CI	$[-1.565,\ 0.574]$	$[-1.010,\ 0.340]$	$[-0.717,\ 0.518]$
Share of positive sample selection [a]	1 (assumed)		0.626 (0.005)
Share of *ac* [b]	0.186 (0.008)		0.185 (0.008)
Annual total expenditure ($N = 16,868$)			
Bounds	$[-0.079,\ 0.186]$	$[-0.019,\ 0.062]$	$[-0.058,\ 0.244]$
Standard Errors	—	$(0.114,\ 0.105)$	$(0.104,\ 0.102)$
95% CI	$[-0.302,\ 0.386]$	$[-0.206,\ 0.234]$	$[-0.228,\ 0.410]$
Share of positive sample selection [a]	1 (assumed)		0.638 (0.005)
Share of *ac* [b]	0.213 (0.008)		0.211 (0.008)

Notes: Outcomes are measured at 12 months after randomization. All calculations use design weights. **CF** refers to the Chen and Flores (2015) bounds under strong sample selection monotonicity without covariates using half-median-unbiased estimates with 95% confidence intervals following Chernozhukov et al. (2013). CF reports no standard errors as inference relies on the CLR projection. **Sharp-basic** refers to the sharp bounds without covariates under strong sample selection monotonicity. **Sharp-DML** refers to the sharp bounds estimated by DML under weak sample selection monotonicity. Standard errors for the two bounds are in parentheses as $(\widehat{SE}_{\text{lower}}, \widehat{SE}_{\text{upper}})$, and the 95% CI for $\theta_{SLATE} = E[Y_1 - Y_0 \mid ac]$ is calculated using Stoye (2020).

[a] Estimated share of positive sample selection $= E[\mathbb{I}^+(X)]$, the fraction of observations in the positive sample selection class (Sharp-DML only. CF and Sharp-basic assume this fraction equals 1).

[b] Estimated share of always-selected compliers π_{ac}.

confidence intervals are also shorter by roughly 25.8%, 36.9%, and 36.0%, respectively. Moreover, for prescription drugs and total expenditure, the Sharp-basic point estimates exclude large portions of the CF range, although their confidence intervals still include zero.

Second, estimates using Sharp-DML are broadly comparable to CF despite relying only on weak sample selection monotonicity. In terms of identified-set length, Sharp-DML reduces the width relative to CF by about 18.5% for outpatient visits and 57.4% for prescription drugs, while for total expenditure the width is similar (slightly larger by about 14%). For confidence intervals, Sharp-DML delivers shorter intervals across all outcomes, with reductions of about 12.7% for outpatient visits, 42.3% for prescription drugs, and 7.3% for total expenditure. Thus, despite the less restrictive assumptions, Sharp-DML maintains or improves precision relative to CF, highlighting the relevance of both sharpness and efficient use of covariates.

Substantively, however, the preferred Sharp-DML confidence intervals include zero for all three outcomes. Hence, the data do not provide statistically precise evidence of intensive margin effects of Medicaid on outpatient visits, prescription-drug use, or annual medical expenditure for always-selected compliers. Taken together with Finkelstein et al. (2012), this suggests that the positive utilization effects found in the original OHIE analysis appear more consistent with a large role for extensive-margin responses, i.e., increased probability of any use, over increases in utilization conditional on positive use.

E.3. Always-selected Complier Profiling

We now conduct the *ac* profiling analysis as discussed in Section 5. Tables E.2–E.4 compare baseline characteristics of always-selected compliers *ac* and the full sample across the three utilization outcomes. A consistent pattern emerges. On average, for outpatient visits and prescription drugs, *ac* are somewhat older (about 1.7 years), more likely to be White, and substantially less likely to be Hispanic, while gender differences are small. Educational attainment is broadly similar, with a slight shift toward some college among *ac*. Differences in lottery timing indicate somewhat greater representation in earlier waves. The most pronounced and robust difference across all outcomes concerns household composition: *ac* are about 10 percentage points more likely to live alone and correspondingly less likely to be in two-person households. For log total expenditure, these differences are

Table E.2: Oregon Health Insurance Experiment Baseline Covariates: Always-Selected Compliers vs. Full Sample (Outcome: Outpatient Visits)

Covariate	*ac*	Full sample	Difference
Age (in yrs.)	45.47 (0.551)	43.76 (0.093)	1.711 (0.545)***
Female	0.630 (0.022)	0.600 (0.004)	0.030 (0.022)
White, non-Hispanic	0.866 (0.016)	0.827 (0.003)	0.039 (0.016)**
Black, non-Hispanic	0.025 (0.009)	0.033 (0.001)	−0.008 (0.009)
Hispanic	0.063 (0.014)	0.114 (0.002)	−0.051 (0.014)***
American Indian/Alaska Native	0.052 (0.012)	0.060 (0.002)	−0.008 (0.011)
Asian	0.026 (0.007)	0.033 (0.001)	−0.008 (0.006)
Other race/ethnicity	0.067 (0.013)	0.092 (0.002)	−0.026 (0.013)*
Education:			
No high school diploma	0.148 (0.017)	0.162 (0.003)	−0.013 (0.017)
High school diploma or GED	0.481 (0.023)	0.495 (0.004)	−0.014 (0.023)
Some college or vocational	0.259 (0.020)	0.225 (0.003)	0.033 (0.019)*
Four-year college degree or higher	0.113 (0.014)	0.119 (0.003)	−0.006 (0.014)
Wave of lottery draws:			
1	0.148 (0.015)	0.115 (0.002)	0.034 (0.014)**
2	0.145 (0.015)	0.115 (0.002)	0.030 (0.015)**
3	0.076 (0.015)	0.114 (0.002)	−0.039 (0.015)***
4	0.142 (0.015)	0.140 (0.003)	0.003 (0.015)
5	0.134 (0.016)	0.142 (0.003)	−0.008 (0.016)
6	0.202 (0.018)	0.204 (0.003)	−0.002 (0.018)
7	0.153 (0.017)	0.171 (0.003)	−0.018 (0.017)
Household members listed:			
1	0.799 (0.020)	0.693 (0.004)	0.106 (0.020)***
2	0.200 (0.020)	0.305 (0.004)	−0.105 (0.020)***
3+	0.002 (0.002)	0.003 (0.000)	−0.001 (0.002)

Notes: $N = 22,064$. The table reports estimated means for the always-selected complier (*ac*) subpopulation and for the full sample, along with their difference (*ac* minus full sample). Standard errors are in parentheses. All estimates use 12-month survey design weights and are based on the weak-monotonicity DML specification with GRF learners. Joint χ^2 tests for the null that all differences within a category are jointly zero: Race/ethnicity ($p = 0.008$); Education ($p = 0.534$); Wave of lottery draws ($p = 0.035$); Household members listed ($p < 0.001$). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.3: Oregon Health Insurance Experiment Baseline Covariates: Always-Selected Compliers vs. Full Sample (Outcome: Prescription Drugs)

Covariate	*ac*	Full sample	Difference
Age (in yrs.)	45.47 (0.551)	43.76 (0.093)	1.711 (0.545)***
Female	0.630 (0.022)	0.600 (0.004)	0.030 (0.022)
White, non-Hispanic	0.866 (0.016)	0.827 (0.003)	0.039 (0.016)**
Black, non-Hispanic	0.025 (0.009)	0.033 (0.001)	−0.008 (0.009)
Hispanic	0.063 (0.014)	0.114 (0.002)	−0.051 (0.014)***
American Indian/Alaska Native	0.052 (0.012)	0.060 (0.002)	−0.008 (0.011)
Asian	0.026 (0.007)	0.033 (0.001)	−0.008 (0.006)
Other race/ethnicity	0.067 (0.013)	0.092 (0.002)	−0.026 (0.013)*
Education:			
No high school diploma	0.148 (0.017)	0.162 (0.003)	−0.013 (0.017)
High school diploma or GED	0.481 (0.023)	0.495 (0.004)	−0.014 (0.023)
Some college or vocational	0.259 (0.020)	0.225 (0.003)	0.033 (0.019)*
Four-year college degree or higher	0.113 (0.014)	0.119 (0.003)	−0.006 (0.014)
Wave of lottery draws:			
1	0.148 (0.015)	0.115 (0.002)	0.034 (0.014)**
2	0.145 (0.015)	0.115 (0.002)	0.030 (0.015)**
3	0.076 (0.015)	0.114 (0.002)	−0.039 (0.015)***
4	0.142 (0.015)	0.140 (0.003)	0.003 (0.015)
5	0.134 (0.016)	0.142 (0.003)	−0.008 (0.016)
6	0.202 (0.018)	0.204 (0.003)	−0.002 (0.018)
7	0.153 (0.017)	0.171 (0.003)	−0.018 (0.017)
Household members listed:			
1	0.799 (0.020)	0.693 (0.004)	0.106 (0.020)***
2	0.200 (0.020)	0.305 (0.004)	−0.105 (0.020)***
3+	0.002 (0.002)	0.003 (0.000)	−0.001 (0.002)

Notes: $N = 17,223$. The table reports estimated means for the always-selected complier (*ac*) subpopulation and for the full sample, along with their difference (*ac* minus full sample). Standard errors are in parentheses. All estimates use 12-month survey design weights and are based on the weak-monotonicity DML specification with GRF learners. Joint χ^2 tests for the null that all differences within a category are jointly zero: Race/ethnicity ($p = 0.008$); Education ($p = 0.534$); Wave of lottery draws ($p = 0.035$); Household members listed ($p = 0.000$). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table E.4: Oregon Health Insurance Experiment Baseline Covariates: Always-Selected Compliers vs. Full Sample (Outcome: Log Annual Total Expenditure)

Covariate	*ac*	Full sample	Difference
Age (in yrs.)	44.06 (0.498)	43.72 (0.094)	0.337 (0.491)
Female	0.607 (0.020)	0.600 (0.004)	0.007 (0.020)
White, non-Hispanic	0.849 (0.016)	0.828 (0.003)	0.021 (0.015)
Black, non-Hispanic	0.028 (0.008)	0.033 (0.001)	−0.005 (0.008)
Hispanic	0.082 (0.013)	0.113 (0.002)	−0.032 (0.013)**
American Indian/Alaska Native	0.062 (0.011)	0.060 (0.002)	0.002 (0.010)
Asian	0.034 (0.006)	0.033 (0.001)	0.000 (0.006)
Other race/ethnicity	0.082 (0.012)	0.092 (0.002)	−0.010 (0.012)
Education:			
No high school diploma	0.155 (0.015)	0.161 (0.003)	−0.006 (0.015)
High school diploma or GED	0.472 (0.020)	0.495 (0.004)	−0.022 (0.020)
Some college or vocational	0.262 (0.017)	0.226 (0.003)	0.036 (0.017)**
Four-year college degree or higher	0.110 (0.013)	0.119 (0.003)	−0.008 (0.012)
Wave of lottery draws:			
1	0.138 (0.013)	0.114 (0.002)	0.024 (0.012)*
2	0.148 (0.013)	0.114 (0.002)	0.034 (0.013)***
3	0.076 (0.014)	0.114 (0.002)	−0.038 (0.014)***
4	0.138 (0.014)	0.141 (0.003)	−0.002 (0.014)
5	0.127 (0.014)	0.142 (0.003)	−0.014 (0.014)
6	0.203 (0.016)	0.205 (0.003)	−0.002 (0.016)
7	0.169 (0.016)	0.170 (0.003)	−0.002 (0.016)
Household members listed:			
1	0.794 (0.018)	0.692 (0.004)	0.102 (0.018)***
2	0.206 (0.018)	0.305 (0.004)	−0.099 (0.018)***
3+	0.000 (0.002)	0.003 (0.000)	−0.002 (0.002)

Notes: $N = 16,868$. The table reports estimated means for the always-selected complier (*ac*) subpopulation and for the full sample, along with their difference (*ac* minus full sample). Standard errors are in parentheses. All estimates use 12-month survey design weights and are based on the weak-monotonicity DML specification with GRF learners. Joint χ^2 tests for the null that all differences within a category are jointly zero: Race/ethnicity ($p = 0.241$); Education ($p = 0.299$); Wave of lottery draws ($p = 0.019$); Household members listed ($p < 0.001$). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

attenuated–age and most demographics are very similar across groups–while household structure and lottery-wave composition remain the main margins of divergence.

Taken together, these patterns indicate mild but systematic selection into the always-selected complier group, primarily along age (for utilization outcomes) and living arrangements. These characteristics are plausibly related to baseline healthcare use conditional on positive utilization. For example, older individuals and single-person households may have more stable or regular interaction with the healthcare system. At the same time, the absence of strong differences in education and other covariates within an already low-income and previously uninsured population suggests that selection is somewhat limited in scope. Given the imprecision of the estimated intensive margin effects, these observable differences do not point to a clear direction for extrapolating θ_{SLATE} to the full sample. If anything, they suggest modest heterogeneity in baseline utilization intensity, but not a strong or systematic gradient that would allow one to infer whether intensive margin responses should be larger or smaller outside the *ac* subpopulation.

F. Supplementary Material for JC and OHIE

F.1. Implementation Details

We summarize the implementation of the DML procedure used in both applications. All nuisance functions are estimated using generalized random forests (Athey et al., 2019), as implemented in the `grf` package in `R`. Propensity scores for the instrument and the joint probabilities of treatment and selection status are estimated via probability forests, while conditional mean outcomes are estimated via regression forests. The conditional distribution function of the outcome, which is required for the trimming quantiles entering the bounds, is estimated by applying probability forests pointwise on a fine grid of trimming values with step size 0.005, and the resulting estimates are post-processed by isotonic regression to enforce monotonicity. For all forests, the minimum leaf size is set to $\lfloor n^{0.65} \rfloor$ and each forest comprises 1,000 trees. Under weak monotonicity, where covariate adjustment is performed, nuisance estimates are obtained via $K = 5$-fold cross-fitting so that the predicted nuisances for each observation are generated by a model trained on the complementary folds, thereby avoiding overfitting bias in the moment conditions. Under strong sample selection monotonicity, no covariates are included, so the nuisance parameters reduce to simple unconditional averages over the whole sample and cross-fitting is not required.

F.2. Balancing Tables for Original Assignment

F.2.1. Job Corps

Because the proposed bounds in Section 3 allow the direction of sample selection to vary with predetermined covariates, we exploit the 27 baseline characteristics available in the Chen and Flores (2015) data. The same set of covariates is used in Lee (2009), and also explored partially in Semenova (2025). These covariates include gender, age at application, months employed in the previous year, usual weekly hours and weekly earnings at the most recent job, race and ethnicity indicators, indicators for children and marital status, parents' education, number of children, highest grade completed, employment at random assignment and in the year before random assignment, lagged earnings, household and personal income categories, and an indicator for whether the person had ever been arrested. Following Schochet et al. (2008), we use design weights throughout because some subpopulations were randomized into the program group with differing, but known, probabilities for programmatic reasons. Table F.1 reports weighted summary statistics for the groups $Z = 0$ and $Z = 1$ as well as their mean differences. The covariates are broadly balanced, as expected under randomization: only age and father's education completed are marginally significant at the 10% level, and usual weekly hours at the most recent job is significant at the 5% level.

F.2.2. Oregon Health Insurance Experiment

We use the publicly available survey data collected and analyzed by Finkelstein et al. (2012). The survey was administered by mail in seven waves during July and August 2009 and includes 29,589 individuals from lottery-selected households and 28,816 individuals from non-selected households. The data also contain pre-randomization demographic information collected at the time of lottery sign-up. Covariates X include gender, age in years, race and ethnicity indicators, education categories, survey-wave fixed effects and household-size fixed effects. Table F.2 reports weighted summary statistics for the $Z = 0$ and $Z = 1$ groups, along with their mean differences, using survey weights. The analysis

Table F.1: Baseline Covariates: Job Corps

Covariate	$Z = 0$	$Z = 1$	Diff. (Std. Err.)
Female	0.458	0.454	−0.004 (0.011)
Age (in yrs.) at baseline	18.35	18.44	0.087 (0.046)*
Black, non-Hispanic	0.491	0.494	0.003 (0.011)
Hispanic	0.172	0.169	−0.003 (0.008)
Other race/ethnicity	0.074	0.072	−0.002 (0.006)
Married	0.023	0.020	−0.003 (0.003)
Living together	0.040	0.039	−0.002 (0.004)
Separated	0.021	0.024	0.003 (0.003)
Has Children	0.193	0.189	−0.004 (0.008)
Number of children	0.268	0.270	0.002 (0.014)
Education	10.11	10.12	0.013 (0.034)
Mother's education	11.47	11.49	0.024 (0.051)
Father's education	11.50	11.41	−0.089 (0.048)*
Ever arrested	0.249	0.248	−0.001 (0.009)
Household income:			
$[\$3,000, \$6,000)$	0.208	0.206	−0.002 (0.007)
$[\$6,000, \$9,000)$	0.114	0.116	0.002 (0.006)
$[\$9,000, \$18,000)$	0.245	0.245	0.001 (0.008)
$\geq \$18,000$	0.181	0.179	−0.001 (0.007)
Personal income:			
$[\$3,000, \$6,000)$	0.131	0.128	−0.003 (0.007)
$[\$6,000, \$9,000)$	0.046	0.053	0.006 (0.004)
$\geq \$9,000$	0.034	0.032	−0.002 (0.004)
At baseline:			
Has job	0.192	0.198	0.007 (0.009)
Months worked, previous year	3.530	3.603	0.074 (0.093)
Had a job, previous year	0.627	0.635	0.008 (0.010)
Earnings, previous year	2815	2909	94.07 (111.7)
Weekly hours, most recent job	20.91	21.83	0.922 (0.454)**
Weekly earnings, most recent job	102.9	111.1	8.183 (5.241)
N	3,599	5,491	

Notes: The table reports sample means and mean differences. All statistics use design weights. The p-value for the joint test that all covariate coefficients in a logit regression of Z on the full set of covariates are zero is 0.549. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

sample size varies by outcome and ranges from 16,868 to 22,064 because of outcome-specific item nonresponse and restrictions to observations with nonmissing treatment, instrument, outcome, and covariates.

Table F.2: Baseline Covariates: Oregon Health Insurance Experiment

Covariate	$Z = 0$	$Z = 1$	Diff. (Std. Err.)
Age (in yrs.)	42.57	42.69	0.087 (0.178)
Female	0.596	0.583	−0.012 (0.007)
White, non-Hispanic	0.827	0.815	−0.012 (0.006)**
Black, non-Hispanic	0.038	0.032	−0.005 (0.003)*
Hispanic	0.123	0.129	0.007 (0.005)
American Indian/Alaska Native	0.067	0.064	−0.003 (0.004)
Asian	0.026	0.030	0.003 (0.002)
Other race/ethnicity	0.103	0.108	0.005 (0.005)
Education:			
No high school diploma	0.175	0.171	−0.005 (0.006)
High school diploma or GED	0.492	0.499	0.007 (0.007)
Some college or vocational	0.220	0.223	0.003 (0.006)
Four-year college degree or higher	0.113	0.107	−0.005 (0.005)
Wave of lottery draws:			
1	0.082	0.146	0.064 (0.005)***
2	0.081	0.152	0.070 (0.005)***
3	0.074	0.153	0.078 (0.005)***
4	0.148	0.133	−0.015 (0.005)***
5	0.151	0.132	−0.018 (0.005)***
6	0.241	0.166	−0.075 (0.006)***
7	0.224	0.117	−0.105 (0.005)***
Household members listed:			
1	0.744	0.659	−0.084 (0.007)***
2	0.255	0.336	0.080 (0.007)***
3+	0.001	0.005	0.004 (0.001)***
N	11,203	11,107	

Notes: The table reports sample means and mean differences. All statistics use 12-month survey design weights. Wave of lottery draws refers to the wave number of the lottery draw in which the household's name(s) appeared on the waitlist. Randomization was within each wave. Household members listed is the number of household members entered on the lottery waitlist. The p-value for the joint test that the coefficients on Age, Female, Race/Ethnicity, and Education are jointly zero in a logit regression of Z on these covariates, lottery-wave fixed effects, and household-size fixed effects is 0.575. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.