Internet Appendix for
"Bounds on Treatment Effects in the Presence of Sample
Selection and Noncompliance: The Wage Effects of Job Corps"

Xuan Chen

School of Labor and Human Resources, Renmin University of China
xchen11@ruc.edu.cn

Carlos A. Flores

Department of Economics, California Polytechnic State University at San Luis Obispo
cflore32@calpoly.edu

January, 2014

Abstract

This internet appendix provides proofs of the two propositions in the paper, a description of the estimation and inference procedure employed, as well as supporting material for the empirical analysis of the wage effects of Job Corps and the simulation study.

1 Proof of Proposition 1

First, we show that under Assumptions 1 through 5 L_{cEE} and U_{cEE} are the smallest and largest possible values, respectively, for the average treatment effect for the stratum cEE . Next, we prove that for $\forall \Delta \in [L_{cEE}, U_{cEE}]$, there exist distributions for cEE , aEE , and cNE consistent with the observed data of Y in $\{Z = 1, D = 1, S = 1\}$ and the constraint that $E[Y^*(1)|aEE] = \bar{Y}^{011}$. In other words, the interval $[L_{cEE}, U_{cEE}]$ contains any other bounds that are consistent with Assumptions 1 through 5. The first-step proof is similar to that in Horowitz and Manski (1995), except that a binding constraint should be satisfied under the lower and upper bounds. Since both L_{cEE} and U_{cEE} depend on the range of $\bar{Y}^{011}$, we need to discuss multiple cases in the proof.

Proof. First by Assumptions 1 through 5, the proportions of each stratum are uniquely determined by the observed data. Thus, the proof is completed given the proportions of the strata. Second since $E[Y^*(0)|cEE]$ is point identified by Assumptions 1 through 5, the proof can be completed with respect to $E[Y^*(1)|cEE]$ instead of the average treatment effect for cEE . Let $\theta = E[Y^*(1)|cEE]$, and then $\Delta = \theta - E[Y^*(0)|cEE]$. Third, since both $LY_{1,cEE}$ and $UY_{1,cEE}$ depend on the range of $\bar{Y}^{011}$, we have to discuss multiple cases.

Let Qy be the observed distribution of Y in the cell $[Z = 1, D = 1, S = 1]$. Let $\alpha_{aEE} = \frac{\pi_{aEE}}{p_{11|1}} = \frac{p_{11|0}}{p_{11|1}}$, $\bar{y}_{aEE} = \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$, and $\tilde{y}_{aEE} = \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$. Denote the expression of $LY_{1,cEE}$ as $\widetilde{LY}_{1,cEE}$ when $\bar{y}_{aEE} > \bar{Y}^{011}$ and the expression of $UY_{1,cEE}$ as $\widetilde{UY}_{1,cEE}$ when $\tilde{y}_{aEE} < \bar{Y}^{011}$. The probability density functions for each stratum are fy_{cEE} , fy_{aEE} , and fy_{cNE} , and their corresponding distributions are Fy_{cEE} , Fy_{aEE} , and Fy_{cNE} .

For the first-step proof, we discuss the two cases to show that $LY_{1,cEE}$ is the smallest possible value for $E[Y^*(1)|cEE]$. The other two cases for the $UY_{1,cEE}$ can be shown in the same way.

1. $LY_{1,cEE} = \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111})$, when $\bar{Y}^{011} \geq \bar{y}_{aEE}$.

Let $Gy = \begin{cases} \frac{Qy}{\alpha_{cEE}}, & \text{if } y \leq y_{\alpha_{cEE}}^{111} \\ 1, & \text{if } y > y_{\alpha_{cEE}}^{111} \end{cases}$.

To show that $\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111})$ is the smallest value, we have $Gy \geq Fy_{cEE}$, for all $Fy_{cEE} \in \{\alpha_{cEE}Fy_{cEE} + \alpha_{aEE}Fy_{aEE} + \alpha_{cNE}Fy_{cNE} = Qy, E[Y^*(1)|aEE] = \bar{Y}^{011}\}$ and all $y \in \mathcal{R}$.

If $y \leq y_{\alpha_{cEE}}^{111}$, $Gy < Fy_{cEE} \Rightarrow Qy < \alpha_{cEE}Fy_{cEE} \Rightarrow Qy < \alpha_{cEE}Fy_{cEE} + \alpha_{aEE}Fy_{aEE} +$

$\alpha_{cNE}Fy_{cNE}$. This contradicts the feasible set of Fy_{cEE} .

If $y > y_{\alpha_{cEE}}^{111}$, $Gy - Fy_{cEE} = 1 - Fy_{cEE} \geq 0$.

2. $LY_{1,cEE} = \widetilde{LY}_{1,cEE}$, when $\bar{Y}^{011} \leq \bar{y}_{aEE}$.

In this case, we first prove that $\bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111})$ is the smallest feasible value for the quantity $\frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}E[Y^*(1)|cEE] + \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}E[Y^*(1)|aEE]$, and then show $\widetilde{LY}_{1,cEE}$ is the smallest feasible value for $E[Y^*(1)|cEE]$.

$$\text{Let } Gy = \begin{cases} \frac{Qy}{1-\alpha_{cNE}}, & \text{if } y \leq y_{1-\alpha_{cNE}}^{111} \\ 1, & \text{if } y > y_{1-\alpha_{cNE}}^{111} \end{cases}.$$

To show $\bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111})$ is the smallest value for the quantity $\frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}E[Y^*(1)|cEE] + \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}E[Y^*(1)|aEE]$, we have $Gy \geq \frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{cEE} + \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{aEE}$, for all $Fy_{cEE} \in \{\alpha_{cEE}Fy_{cEE} + \alpha_{aEE}Fy_{aEE} + \alpha_{cNE}Fy_{cNE} = Qy, E[Y^*(1)|aEE] = \bar{Y}^{011}\}$ and all $y \in \mathcal{R}$.

If $y \leq y_{1-\alpha_{cNE}}^{111}$, $Gy < \frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{cEE} + \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{aEE} \Rightarrow Qy < \alpha_{cEE}Fy_{cEE} + \alpha_{aEE}Fy_{aEE} \Rightarrow Qy < \alpha_{cEE}Fy_{cEE} + \alpha_{aEE}Fy_{aEE} + \alpha_{cNE}Fy_{cNE}$. This contradicts the feasible set of Fy_{cEE} .

If $y > y_{1-\alpha_{cNE}}^{111}$, we have $Gy - (\frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{cEE} + \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{aEE}) = 1 - \frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{cEE} - \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}Fy_{aEE} \geq 0$.

Since $E[Y^*(1)|aEE] = \bar{Y}^{011}$, $\widetilde{LY}_{1,cEE}$ is the smallest value for $E[Y^*(1)|cEE]$.

For $UY_{1,cEE}$, $UY_{1,cEE} = \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})$ is obtained by considering the observations in the top distribution with the proportion $\frac{\pi_{cEE}}{p_{11|1}}$ as members of cEE . This unconstrained solution has to satisfy the constraint that $\bar{Y}^{011} \leq \bar{y}_{aEE}$, which equals $\bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$. Otherwise, $UY_{1,cEE}$ is derived by the equation $\bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) = UY_{1,cEE} \frac{\pi_{cEE}}{\pi_{cEE}+\pi_{aEE}} + \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cEE}+\pi_{aEE}}$. Thus, when $\bar{Y}^{011} \geq \bar{y}_{aEE}$, $UY_{1,cEE} = \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) \frac{q_{1|0}-p_{01|1}}{p_{01|0}-p_{01|1}} - \bar{Y}^{011} \frac{p_{11|0}}{p_{01|0}-p_{01|1}} = \widetilde{UY}_{1,cEE}$. The proof that $UY_{1,cEE}$ is the largest value for the upper bound of $E[Y^*(1)|cEE]$ is similar to that for the lower bound.

For the second-step proof, we have four cases to discuss, taking into account the lower and upper bounds simultaneously. Since they form different segmentations of Qy , in each case we use some cutoff values to discuss $\forall \theta \in [LY_{1,cEE}, cutoff]$ and $\forall \theta \in [cutoff, UY_{1,cEE}]$ separately. In either interval for θ , we have to discuss the range of $\bar{Y}^{011}$ as we have done in the first-step proof. The four cases are listed as follows.

$$1. LY_{1,cEE} = \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), UY_{1,cEE} = \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})$$

It happens when $\bar{y}_{aEE} \leq \bar{Y}^{011} \leq \tilde{y}_{aEE}$; in other words, $\pi_{cNE} \geq \pi_{cEE}$. Since $\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}) \leq \bar{Y}^{111} \leq \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})$, $cutoff = \bar{Y}^{111}$.

For $\forall \theta \in [\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), cutoff]$, two cases to discuss are $\bar{Y}^{011} \geq \bar{y}_{aEE}$ & $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\bar{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

For $\forall \theta \in [cutoff, \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})]$, two cases to discuss are $\bar{Y}^{011} \leq \tilde{y}_{aEE}$ & $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$ and $\bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111}) \leq \bar{Y}^{011} \leq \tilde{y}_{aEE}$.

$$2. LY_{1,cEE} = \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), UY_{1,cEE} = \widetilde{UY}_{1,cEE}$$

It happens when $\bar{Y}^{011} \geq \max\{\bar{y}_{aEE}, \tilde{y}_{aEE}\}$. $LY_{1,cEE} = \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}) \leq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\widetilde{UY}_{1,cEE} = \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cEE}} \geq \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111}) \frac{\pi_{aEE}}{\pi_{cEE}} = \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$. The inequality for $\widetilde{UY}_{1,cEE}$ is derived from $\bar{Y}^{011} \leq \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111})$, and the equality is derived from the equation $\bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) = \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{aEE}} \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}) + \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111})$. Thus, $\widetilde{UY}_{1,cEE} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}) \geq \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111})$. Therefore, $cutoff = \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

For $\forall \theta \in [\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), cutoff]$, two cases to discuss are $\bar{Y}^{011} \geq \bar{y}_{aEE}$ & $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\bar{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

For $\forall \theta \in [cutoff, \widetilde{UY}_{1,cEE}]$, two cases to discuss are $\bar{Y}^{011} \geq \tilde{y}_{aEE}$ & $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\tilde{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

$$3. LY_{1,cEE} = \widetilde{LY}_{1,cEE}, UY_{1,cEE} = \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})$$

It happens when $\bar{Y}^{011} \leq \min\{\bar{y}_{aEE}, \tilde{y}_{aEE}\}$. $UY_{1,cEE} = \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111}) \geq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$ and $\widetilde{LY}_{1,cEE} = \bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111}) \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cEE}} \leq \bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111}) \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - \bar{Y}(Y \leq y_{\alpha_{aEE}}^{111}) \frac{\pi_{aEE}}{\pi_{cEE}} = \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$. The inequality for $\widetilde{LY}_{1,cEE}$ is derived from $\bar{Y}^{011} \geq \bar{Y}(Y \leq y_{\alpha_{aEE}}^{111})$, and the equality is derived from the equation $\bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111}) = \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{aEE}} \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111}) + \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} \bar{Y}(Y \leq y_{\alpha_{aEE}}^{111})$. Thus, $\widetilde{LY}_{1,cEE} \geq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111}) \geq \widetilde{LY}_{1,cEE}$. Therefore, $cutoff = \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$.

For $\forall \theta \in [\widetilde{LY}_{1,cEE}, cutoff]$, two cases to discuss are $\bar{Y}^{011} \leq \bar{y}_{aEE}$ & $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$ and $\bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111}) \leq \bar{Y}^{011} \leq \bar{y}_{aEE}$.

For $\forall \theta \in [cutoff, \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})]$, two cases to discuss are $\bar{Y}^{011} \leq \tilde{y}_{aEE}$ & $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$ and $\bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111}) \leq \bar{Y}^{011} \leq \tilde{y}_{aEE}$.

$$4. LY_{1,cEE} = \widetilde{LY}_{1,cEE}, UY_{1,cEE} = \widetilde{UY}_{1,cEE}$$

It happens when $\widetilde{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{y}_{aEE}$; in other words, $\pi_{cNE} \leq \pi_{cEE}$. In this case, it is difficult to get a uniform cutoff. We discuss multiple cases conditional on the proportions of the strata.

When $\pi_{cNE} \leq \pi_{cEE}$ and $\pi_{aEE} \leq \pi_{cEE}$, we discuss the intervals $\forall \theta \in [\widetilde{LY}_{1,cEE}, \bar{y}_{aEE}]$ and $\forall \theta \in [\widetilde{y}_{aEE}, \widetilde{UY}_{1,cEE}]$ to complete the proof in the entire range, i.e., $\forall \theta \in [\widetilde{LY}_{1,cEE}, \widetilde{UY}_{1,cEE}]$.

For $\forall \theta \in [\widetilde{LY}_{1,cEE}, \bar{y}_{aEE}]$, the two cases are $\bar{Y}^{011} \leq \bar{y}_{aEE}$ & $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$ and $\bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111}) \leq \bar{Y}^{011} \leq \bar{y}_{aEE}$.

For $\forall \theta \in [\widetilde{y}_{aEE}, \widetilde{UY}_{1,cEE}]$, the two cases are $\bar{Y}^{011} \geq \widetilde{y}_{aEE}$ & $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\widetilde{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

When $\pi_{cNE} \leq \pi_{cEE} \leq \pi_{aEE}$, we discuss the intervals $\forall \theta \in [\widetilde{LY}_{1,cEE}, \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})]$ and $\forall \theta \in [\bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}), \widetilde{UY}_{1,cEE}]$ to complete the proof in the entire range.

For $\forall \theta \in [\widetilde{LY}_{1,cEE}, \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})]$, we have $\bar{Y}^{011} \leq \bar{y}_{aEE}$ & $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$.

For $\forall \theta \in [\bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}), \widetilde{UY}_{1,cEE}]$, we have $\bar{Y}^{011} \geq \widetilde{y}_{aEE}$ & $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

From all the above, we can find that though cutoffs are different in the four cases, the discussion of $\bar{Y}^{011}$ is repeated. Cases 2 and 3 compose a complete discussion of $\bar{Y}^{011}$. In the following, we only write the proof of Case 2. Case 3 can be shown in a similar way.

$$2. LY_{1,cEE} = \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), UY_{1,cEE} = \widetilde{UY}_{1,cEE}, \text{ and } cutoff = \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}).$$

First, $\forall \theta \in [cutoff, \widetilde{UY}_{1,cEE}]$, $\exists \lambda \in (0, 1]$, s.t. $\lambda \widetilde{UY}_{1,cEE} + (1 - \lambda) \bar{Y}(Y \leq y_{\alpha_{cNE}}^{111}) = \theta$, since $\bar{Y}(Y \leq y_{\alpha_{cNE}}^{111}) \leq \theta \leq \widetilde{UY}_{1,cEE}$. ($\lambda = \frac{\theta - \bar{Y}(Y \leq y_{\alpha_{cNE}}^{111})}{\widetilde{UY}_{1,cEE} - \bar{Y}(Y \leq y_{\alpha_{cNE}}^{111})}$.)

To construct $f_{y_{aEE}}$, it is necessary to discuss the value of $\bar{Y}^{011}$. One case is $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$. This happens when either $\pi_{cEE} \leq \pi_{aEE}$, or $\pi_{cEE} \geq \pi_{aEE}$ but aEE take up the very top quantiles of the observed distribution. The other case is $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$. This is true only when $\pi_{cEE} \geq \pi_{aEE}$.

$$(1) \bar{Y}^{011} \geq \widetilde{y}_{aEE} \text{ \& } \bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$$

Let ty be the observed density of $Y(y \geq y_{1-\alpha_{aEE}}^{111})$, hy the density of $Y(y_{\alpha_{cNE}}^{111} \leq Y \leq$

$y_{1-\alpha_{aEE}}^{111}$) and gy the density of $Y(y \leq y_{\alpha_{cNE}}^{111})$. Since $\bar{Y}(Y \leq y_{1-\alpha_{aEE}}^{111}) \leq \bar{Y}^{011} \leq \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111})$, $\exists \tau \in (0, 1]$, s.t. $\tau \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111}) + (1 - \tau) \bar{Y}(Y \leq y_{1-\alpha_{aEE}}^{111}) = \bar{Y}^{011}$. $\tau ty + (1 - \tau) \frac{\pi_{cEE}}{\pi_{cNE} + \pi_{cEE}} hy + (1 - \tau) \frac{\pi_{cNE}}{\pi_{cNE} + \pi_{cEE}} gy = fy_{aEE}$. To construct f_{cEE} , we use the equation $\lambda \widetilde{UY}_{1,cEE} + (1 - \lambda) \bar{Y}(Y \leq y_{\alpha_{cNE}}^{111}) = \theta$. Since $\widetilde{UY}_{1,cEE} = \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cEE}}$ and it is obtained by temporarily assuming values of aEE members are above $y_{\alpha_{cNE}}^{111}$, we have $\exists \phi \in [0, 1]$, s.t. $\phi \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111}) + (1 - \phi) \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111}) = \bar{Y}^{011}$. Then, the density, whose mean equals $\widetilde{UY}_{1,cEE}$, is $[\frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} ty + \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{aEE}} hy] \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}}$. Thus, by the equation to obtain θ , we have $\lambda \{ \frac{\pi_{aEE}}{\pi_{cEE}} ty + hy - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}} \} + (1 - \lambda) gy = fy_{cEE}$. Since $\alpha_{aEE} ty + \alpha_{cEE} hy + \alpha_{cNE} gy = \alpha_{aEE} fy_{aEE} + \alpha_{cEE} fy_{cEE} + \alpha_{cNE} fy_{cNE}$, the corresponding density for cNE is $\frac{\pi_{aEE}}{\pi_{cNE}} [1 - \tau - \lambda(1 - \phi)] ty + \frac{\pi_{cEE}}{\pi_{cNE}} [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}} - \lambda(1 - (1 - \phi) \frac{\pi_{aEE}}{\pi_{cEE}})] hy + [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}} - (1 - \lambda) \frac{\pi_{cEE}}{\pi_{cNE}}] gy = fy_{cNE}$.

$$(2) \tilde{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$$

Let ty be the observed density of $Y(y \geq y_{1-\alpha_{cEE}}^{111})$, hy the density of $Y(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$ and gy the density of $Y(y \leq y_{\alpha_{cNE}}^{111})$. Since $\bar{Y}(Y \leq y_{\alpha_{cNE}}^{111}) \leq \bar{Y}^{011} \leq \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111})$, $\exists \tau \in (0, 1)$, s.t. $\tau \bar{Y}(Y \geq y_{\alpha_{cNE}}^{111}) + (1 - \tau) \bar{Y}(Y \leq y_{\alpha_{cNE}}^{111}) = \bar{Y}^{011}$. $\tau \frac{\pi_{cEE}}{\pi_{aEE} + \pi_{cEE}} ty + \tau \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cEE}} hy + (1 - \tau) gy = fy_{aEE}$. To construct f_{cEE} , as in (1), since $\widetilde{UY}_{1,cEE}$ is obtained by temporarily assuming values of aEE are above $y_{\alpha_{cNE}}^{111}$, $\exists \phi \in [0, 1]$, s.t. $\phi \bar{Y}(Y \geq y_{1-\alpha_{cEE}}^{111}) + (1 - \phi) \tilde{y}_{aEE} = \bar{Y}^{011}$. Then, the density, whose mean equals $\widetilde{UY}_{1,cEE}$, is $[\frac{\pi_{cEE}}{\pi_{cEE} + \pi_{aEE}} ty + \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} hy] \frac{\pi_{cEE} + \pi_{aEE}}{\pi_{cEE}} - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}}$. Thus, $\lambda \{ ty + \frac{\pi_{aEE}}{\pi_{cEE}} hy - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}} \} + (1 - \lambda) gy = fy_{cEE}$. Similarly as in (1), we finally have $\frac{\pi_{cEE}}{\pi_{cNE}} [1 - \tau \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} - \lambda(1 - \phi \frac{\pi_{aEE}}{\pi_{cEE}})] ty + \frac{\pi_{aEE}}{\pi_{cNE}} [1 - \tau \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} - \lambda \phi] hy + [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cNE}} - (1 - \lambda) \frac{\pi_{cEE}}{\pi_{cNE}}] gy = fy_{cNE}$.

Second, $\forall \theta \in [\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}), \text{cutoff}]$, $\exists \lambda \in [0, 1)$, s.t. $\lambda \bar{Y}(Y \geq y_{\alpha_{cEE}}^{111}) + (1 - \lambda) \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}) = \theta$, since $\bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}) \leq \theta \leq \bar{Y}(Y \geq y_{\alpha_{cEE}}^{111})$. ($\lambda = \frac{\theta - \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111})}{\bar{Y}(Y \geq y_{\alpha_{cEE}}^{111}) - \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111})}$.)

To discuss the value of $\bar{Y}^{011}$, one case is $\bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$. It happens when either $\pi_{cNE} \leq \pi_{aEE}$, or $\pi_{cNE} \geq \pi_{aEE}$ but aEE take up very top quantiles of the observed distribution. The other case is $\bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ when $\pi_{cNE} \geq \pi_{aEE}$.

$$(1) \bar{Y}^{011} \geq \bar{y}_{aEE} \ \& \ \bar{Y}^{011} \geq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$$

Let ty be the observed density of $Y(y \geq y_{1-\alpha_{aEE}}^{111})$, hy the density of $Y(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$, and gy the density of $Y(y \leq y_{\alpha_{cEE}}^{111})$. Since $\bar{Y}(Y \leq y_{1-\alpha_{aEE}}^{111}) \leq \bar{Y}^{011} \leq \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111})$,

$y_{1-\alpha_{aEE}}^{111}$), $\exists \tau \in (0, 1]$, s.t. $\tau \bar{Y}(Y \geq y_{1-\alpha_{aEE}}^{111}) + (1 - \tau) \bar{Y}(Y \leq y_{1-\alpha_{aEE}}^{111}) = \bar{Y}^{011}$. $\tau ty + (1 - \tau) \frac{\pi_{cNE}}{\pi_{cNE} + \pi_{cEE}} hy + (1 - \tau) \frac{\pi_{cEE}}{\pi_{cNE} + \pi_{cEE}} gy = fy_{aEE}$. For cEE , by the equation $\lambda \bar{Y}(Y \geq y_{\alpha_{cEE}}^{111}) + (1 - \lambda) \bar{Y}(Y \leq y_{\alpha_{cEE}}^{111}) = \theta$, we have $\lambda (\frac{\pi_{aEE}}{\pi_{cNE} + \pi_{aEE}} ty + \frac{\pi_{cNE}}{\pi_{cNE} + \pi_{aEE}} hy) + (1 - \lambda) gy = fy_{cEE}$. Finally, $\frac{\pi_{aEE}}{\pi_{cNE}} (1 - \tau - \lambda \frac{\pi_{cEE}}{\pi_{aEE} + \pi_{cNE}}) ty + [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}} - \lambda \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}}] hy + \frac{\pi_{cEE}}{\pi_{cNE}} [\lambda - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}}] gy = fy_{cNE}$.

$$(2) \bar{y}_{aEE} \leq \bar{Y}^{011} \leq \bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$$

Let ty be the observed density of $Y(y \geq y_{1-\alpha_{cNE}}^{111})$, hy the density of $Y(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$, and gy the density of $Y(y \leq y_{\alpha_{cEE}}^{111})$. Since $\bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111}) \leq \bar{Y}^{011} \leq \bar{Y}(Y \geq y_{1-\alpha_{cNE}}^{111})$, $\exists \tau \in (0, 1]$, s.t. $\tau \bar{Y}(Y \geq y_{1-\alpha_{cNE}}^{111}) + (1 - \tau) \bar{Y}(Y \leq y_{1-\alpha_{cNE}}^{111}) = \bar{Y}^{011}$. $\tau ty + (1 - \tau) \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cEE}} hy + (1 - \tau) \frac{\pi_{cEE}}{\pi_{aEE} + \pi_{cEE}} gy = fy_{aEE}$. For cEE , by the equation to obtain θ , we have $\lambda (\frac{\pi_{cNE}}{\pi_{cNE} + \pi_{aEE}} ty + \frac{\pi_{aEE}}{\pi_{cNE} + \pi_{aEE}} hy) + (1 - \lambda) gy = fy_{cEE}$. Finally, $(1 - \tau \frac{\pi_{aEE}}{\pi_{cNE}} - \lambda \frac{\pi_{cEE}}{\pi_{aEE} + \pi_{cNE}}) ty + \frac{\pi_{aEE}}{\pi_{cNE}} [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} - \lambda \frac{\pi_{cEE}}{\pi_{aEE} + \pi_{cNE}}] hy + \frac{\pi_{cEE}}{\pi_{cNE}} [\lambda - (1 - \tau) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}}] gy = fy_{cNE}$. ■

2 Proof of Proposition 2

The proof of Proposition 2 is similar to that of Proposition 1, except for two differences: multiple cases reduce to two because there is only one term for the lower bound in Proposition 2, $\bar{L}_{cEE}$ (as opposed to the lower bound in Proposition 1, which is the maximum of two terms), and the constructed distributions for cEE and cNE should also satisfy the mean dominance assumption.

Proof. As in the proof of Proposition 1, we first show that $\bar{LY}_{1,cEE}$ is the smallest feasible value for $E[Y^*(1)|cEE]$, and then show $\forall \theta \in [\bar{LY}_{1,cEE}, UY_{1,cEE}]$, there exist distributions Fy_{cEE} , Fy_{aEE} and Fy_{cNE} satisfying Assumptions 1 through 6.

For $\forall \theta \in [\bar{LY}_{1,cEE}, UY_{1,cEE}]$, by equation (6), we have

$$\begin{aligned} & E[Y^*(1)|cEE] - E[Y^*(1)|cNE] \\ &= \theta - (\bar{Y}^{111} \frac{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}}{\pi_{cNE}} - \theta \frac{\pi_{cEE}}{\pi_{cNE}} - \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cNE}}) \\ &\geq (1 + \frac{\pi_{cEE}}{\pi_{cNE}}) \bar{LY}_{1,cEE} + \bar{Y}^{011} \frac{\pi_{aEE}}{\pi_{cNE}} - \bar{Y}^{111} \frac{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}}{\pi_{cNE}} = 0. \end{aligned}$$

If there was another lower bound smaller than $\bar{LY}_{1,cEE}$, $E[Y^*(1)|cEE] - E[Y^*(1)|cNE]$ would be negative when $E[Y^*(1)|cEE]$ reached that lower bound. This contradicts Assumption 6. Thus, $\bar{LY}_{1,cEE}$ is the smallest value for $E[Y^*(1)|cEE]$ under Assumptions 1 through 6.

Next, similarly as the proof of Proposition 1, let's show that for $\forall \theta \in [\bar{LY}_{1,cEE}, UY_{1,cEE}]$,

there exist distributions Fy_{cEE} , Fy_{aEE} , and Fy_{cNE} . Due to the fixed value of $\overline{LY}_{1,cEE}$, now we only have two general cases depending on the value of $UY_{1,cEE}$.

$$1. LY_{1,cEE} = \overline{LY}_{1,cEE}, UY_{1,cEE} = \overline{Y}(Y \geq y_{1-\alpha_{cEE}}^{111}) \text{ when } \overline{Y}^{011} \leq \tilde{y}_{aEE}$$

As in the proof of Proposition 1, we have to discuss: $\overline{Y}^{011} \leq \tilde{y}_{aEE}$ & $\overline{Y}^{011} \leq \overline{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$ and $\overline{Y}(y_{\alpha_{aEE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111}) \leq \overline{Y}^{011} \leq \tilde{y}_{aEE}$.

$$2. LY_{1,cEE} = \overline{LY}_{1,cEE}, UY_{1,cEE} = \widetilde{UY}_{1,cEE} \text{ when } \overline{Y}^{011} \geq \tilde{y}_{aEE}$$

Correspondingly, we have to discuss: $\overline{Y}^{011} \geq \tilde{y}_{aEE}$ & $\overline{Y}^{011} \geq \overline{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ and $\tilde{y}_{aEE} \leq \overline{Y}^{011} \leq \overline{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$.

The proof of the four cases above is similar to the second-step proof of Proposition 1. In the following, we take Case $\tilde{y}_{aEE} \leq \overline{Y}^{011} \leq \overline{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$ as an example to illustrate the second-step proof of Proposition 2.

$$\text{Case: } \tilde{y}_{aEE} \leq \overline{Y}^{011} \leq \overline{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$$

When $\tilde{y}_{aEE} \leq \overline{Y}^{011}$, under Assumptions 1 through 6 we have $LY_{1,cEE} = \overline{LY}_{1,cEE}$ and $UY_{1,cEE} = \widetilde{UY}_{1,cEE}$. Let ty be the observed density of $Y(y \geq y_{1-\alpha_{cEE}}^{111})$, hy the density of $Y(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$, and gy the density of $Y(y \leq y_{\alpha_{cNE}}^{111})$.

$$\forall \theta \in [\overline{LY}_{1,cEE}, \widetilde{UY}_{1,cEE}], \exists \lambda \in [0, 1], \text{ s.t. } \lambda \widetilde{UY}_{1,cEE} + (1 - \lambda) \overline{LY}_{1,cEE} = \theta.$$

To construct fy_{aEE} , since $\tilde{y}_{aEE} \leq \overline{Y}^{011} \leq \overline{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{aEE}}^{111})$, $\exists \tau \in (0, 1]$, s.t. $\tau \overline{Y}(Y \geq y_{1-\alpha_{cEE}}^{111}) + (1 - \tau) \overline{Y}(Y \leq y_{1-\alpha_{cEE}}^{111}) = \overline{Y}^{011}$. Then, $\tau ty + (1 - \tau) \frac{\pi_{aEE}}{\pi_{cNE} + \pi_{aEE}} hy + (1 - \tau) \frac{\pi_{cNE}}{\pi_{cNE} + \pi_{aEE}} gy = fy_{aEE}$. For fy_{cEE} , again we first construct the densities whose means equal $\widetilde{UY}_{1,cEE}$ and $\overline{LY}_{1,cEE}$, respectively. Since $\widetilde{UY}_{1,cEE}$ is obtained by temporarily assuming values of aEE are above $y_{\alpha_{cNE}}^{111}$, $\exists \phi \in [0, 1]$, s.t. $\phi \overline{Y}(Y \geq y_{1-\alpha_{cEE}}^{111}) + (1 - \phi) \tilde{y}_{aEE} = \overline{Y}^{011}$. Then, the density for $\widetilde{UY}_{1,cEE}$ is $ty + \frac{\pi_{aEE}}{\pi_{cEE}} hy - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}}$. Since $\overline{LY}_{1,cEE} = \frac{p_{11|1} \overline{Y}^{111} - \pi_{aEE} \overline{Y}^{011}}{\pi_{cEE} + \pi_{cNE}}$, its corresponding density equals $(\pi_{cEE} ty + \pi_{aEE} hy + \pi_{cNE} gy - \pi_{aEE} fy_{aEE}) / (\pi_{cEE} + \pi_{cNE})$. By the equation $\lambda \widetilde{UY}_{1,cEE} + (1 - \lambda) \overline{LY}_{1,cEE} = \theta$, we have $\lambda \{ty + \frac{\pi_{aEE}}{\pi_{cEE}} hy - [\phi ty + (1 - \phi) hy] \frac{\pi_{aEE}}{\pi_{cEE}}\} + (1 - \lambda) (\frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}} ty + \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}} hy + \frac{\pi_{cNE}}{\pi_{cEE} + \pi_{cNE}} gy - fy_{aEE} \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}}) = fy_{cEE}$. After plugging fy_{aEE} into the equation for fy_{cEE} , $[\lambda(1 - \phi \frac{\pi_{aEE}}{\pi_{cEE}}) + (1 - \lambda)(\frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}} - \tau \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}})] ty + \{\lambda \phi \frac{\pi_{aEE}}{\pi_{cEE}} + (1 - \lambda) \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}} [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cNE}}]\} hy + (1 - \lambda) \frac{\pi_{cNE}}{\pi_{cEE} + \pi_{cNE}} [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cNE}}] gy = fy_{cEE}$. Then, the corresponding density for cNE is $\{-\frac{\pi_{aEE}}{\pi_{cNE}} \tau + \frac{\pi_{cEE}}{\pi_{cNE}} [1 - \lambda(1 - \phi \frac{\pi_{aEE}}{\pi_{cEE}}) - (1 - \lambda)(\frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}} - \tau \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{cNE}})]\} ty + \frac{\pi_{aEE}}{\pi_{cNE}} \{-\lambda \phi - [1 - (1 - \lambda) \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}}]\} [1 -$

$$(1 - \tau) \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cNE}}] \} hy + [1 - (1 - \lambda) \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{cNE}}] [1 - (1 - \tau) \frac{\pi_{aEE}}{\pi_{aEE} + \pi_{cNE}}] gy = fy_{cNE}.$$

The inequality in the second paragraph of this proof shows that Assumption 6 holds, as long as $E[Y^*(1)|cEE] \geq \overline{LY}_{1,cEE}$. Since $\theta \in [\overline{LY}_{1,cEE}, UY_{1,cEE}]$ holds by construction, the constructed densities satisfy Assumption 6. ■

3 Procedure for Estimation and Inference

As discussed in the paper, Chernozhukov, Lee, and Rosen (2013)'s (hereafter CLR) precision-corrected adjustment consists of adding to each estimated bounding function its pointwise standard error times an appropriate critical value, $k(p)$.

The selection of $k(p)$ relies on a standardized Gaussian process $Z_n^*(v)$. For any compact set $V \in \mathcal{V}$, CLR approximate by simulation the p -th quantile of $\sup_{v \in V} Z_n^*(v)$, denoted by $k_{n,V}(p)$, and use it in place of $k(p)$. Since setting $V = \mathcal{V}^u$ for the upper bound leads to asymptotically valid but conservative inference, they propose a preliminary set estimator $\hat{V}_n^u$ of $V_0^u = \arg \min_{v \in \mathcal{V}^u} \theta^u(v)$, which they call an adaptive inequality selector. Intuitively, $\hat{V}_n^u$ selects the bounding functions that are close enough to binding to affect the asymptotic distribution of the estimator of the upper bound. Similarly, a preliminary set estimator $\hat{V}_n^l$ of $V_0^l = \arg \max_{v \in \mathcal{V}^l} \theta^l(v)$ is used for the lower bound.

As an illustration, the upper bound for $\theta_0 = E[Y^*(1)|cEE]$ in Proposition 1 is given by $\theta_0^u = \min_{v \in \mathcal{V}^u = \{1,2\}} \theta^u(v)$, with $\theta^u(1) = \overline{Y}(Y \geq y_{1-\alpha_{cEE}}^{111})$ and $\theta^u(2) = \overline{Y}(Y \geq y_{\alpha_{cNE}}^{111}) \frac{q_{1|0} - p_{01|1}}{p_{01|0} - p_{01|1}} - \overline{Y}^{011} \frac{p_{11|0}}{p_{01|0} - p_{01|1}}$. The precision-corrected estimate of θ_0^u is given by

$$\hat{\theta}^u(p) = \min_{v \in \{1,2\}} [\hat{\theta}^u(v) + k_{n,\hat{V}_n^u}^u(p) s^u(v)], \quad (1)$$

where $\hat{\theta}^u(v)$ is the sample analog estimate of $\theta^u(v)$ and $s^u(v)$ is its standard error. Let $\gamma_n = [\theta_n^u(1) \ \theta_n^u(2)]'$ be the vector containing the two bounding functions and let $\hat{\gamma}_n = [\hat{\theta}_n^u(1) \ \hat{\theta}_n^u(2)]'$ denote its sample analog estimator. The steps to compute the set estimator $\hat{V}_n^u$ and the critical value $k_{n,\hat{V}_n^u}^u(p)$ in (1) are:

(1) We obtain by bootstrapping a consistent estimate $\hat{\Omega}_n$ of the asymptotic variance of $\sqrt{n}(\hat{\gamma}_n - \gamma_n)$. Let $\hat{g}_n(v)'$ denote the v^{th} row of $\hat{\Omega}_n^{1/2}$ and let $s_n^u(v) = \|\hat{g}_n(v)\| / \sqrt{n}$.

(2) We simulate R draws from $\mathcal{N}(0, I_2)$, denoted $Z_1, \dots, Z_R$, where I_2 is a 2×2 identity

matrix, and let $Z_r^*(v) = \hat{g}_n(v)'Z_r / \|\hat{g}_n(v)\|$ for $r = 1, \dots, R$.

(3) Let $Q_p(X)$ denote the p -th quantile of a random variable X and, following CLR, set $c_n = 1 - (.1/\log n)$. We compute $k_{n,\mathcal{V}^u}^u(c_n) = Q_{c_n}(\max_{v \in \{1,2\}} Z_r^*(v), r = 1, \dots, R)$; that is, for each replication r we calculate the maximum of $Z_r^*(1)$ and $Z_r^*(2)$, and take the c -th quantile of those R values. We then use $k_{n,\mathcal{V}^u}^u(c_n)$ to compute $\hat{V}_n^u = \{v \in \{1,2\} : \hat{\theta}^u(v) \leq \min_{\tilde{v} \in \{1,2\}} \{\hat{\theta}^u(\tilde{v}) + k_{n,\mathcal{V}^u}^u(c_n)s_n^u(\tilde{v})\} + 2k_{n,\mathcal{V}^u}^u(c_n)s_n^u(v)\}$.

(4) We compute $k_{n,\hat{V}_n^u}^u(p) = Q_p(\max_{v \in \hat{V}_n^u} Z_r^*(v), r = 1, \dots, R)$, so that the critical value is based on $\hat{V}_n^u$ instead of $\mathcal{V}^u = \{1,2\}$.

The precision-corrected estimate of the lower bound θ_0^l is given by

$$\hat{\theta}^l(p) = \max_{v \in \{1,2\}} [\hat{\theta}^l(v) - k_{n,\hat{V}_n^l}^l(p)s^l(v)], \quad (2)$$

where $\hat{\theta}^l(v)$ is the sample analog estimate of $\theta^l(v)$ and $s^l(v)$ is its standard error. To compute $k_{n,\hat{V}_n^l}^l(p)$ in (2), we follow the same steps above, but in step (3) $\hat{V}_n^u$ is replaced by $\hat{V}_n^l = \{v \in \{1,2\} : \hat{\theta}^l(v) \geq \max_{\tilde{v} \in \{1,2\}} [\hat{\theta}^l(\tilde{v}) - k_{n,\mathcal{V}^l}^l(c_n)s_n^l(\tilde{v})] - 2k_{n,\mathcal{V}^l}^l(c_n)s_n^l(v)\}$. Because of the symmetry of the normal distribution, no changes are needed when computing the quantiles in steps (3) and (4).

Half-median-unbiased estimators of the upper and lower bounds are obtained by setting $p = 1/2$ in the steps above and using equations (1) and (2) to compute $\hat{\theta}^u(1/2)$ and $\hat{\theta}^l(1/2)$, respectively. To construct confidence intervals for the parameter θ_0 , it is important to take into account the length of the identified set. Following CLR, let $\hat{\Gamma}_n = \hat{\theta}_n^u(1/2) - \hat{\theta}_n^l(1/2)$, $\hat{\Gamma}_n^+ = \max(0, \hat{\Gamma}_n)$, $\rho_n = \max\{\hat{\theta}_n^u(3/4) - \hat{\theta}_n^u(1/4), \hat{\theta}_n^l(1/4) - \hat{\theta}_n^l(3/4)\}$, $\tau_n = 1/(\rho_n \log n)$ and $\hat{p}_n = 1 - \Phi(\tau_n \hat{\Gamma}_n^+)$, where $\Phi(\cdot)$ is the standard normal CDF. Note that $\hat{p}_n \in [1 - \alpha, 1 - \alpha/2]$, with $\hat{p}_n = 1 - \alpha/2$ when $\hat{\Gamma}_n = 0$ and $\hat{p}_n$ approaching $1 - \alpha$ when $\hat{\Gamma}_n$ grows large relative to sampling error. Then, an asymptotically valid $1 - \alpha$ confidence interval for θ_0 is given by $[\hat{\theta}_n^l(\hat{p}_n), \hat{\theta}_n^u(\hat{p}_n)]$.

Finally, note that the lower bounds in Proposition 2 do not involve the max operator. However, for consistency, we employ the same methodology described above (where $\mathcal{V}$ is a singleton for the lower bound) to construct confidence intervals and we report the half-median-unbiased estimates of the lower bounds (which, not surprisingly, are practically the same as the

sample analog estimates).

4 Estimation of Average Baseline Characteristics of the Strata

We write the moment functions for average baseline characteristics of all the strata based on the conditional expectation in each cell defined by $\{Z, D, S\}$. Let $\bar{x}_k$ denote the expectation of a scalar baseline variable for a certain stratum k . The moment function for this variable is defined as:

$$g(\{\bar{x}_k\}) = \begin{bmatrix} (x - \bar{x}_{aNN})(1 - Z)D(1 - S) \\ (x - \bar{x}_{aEE})(1 - Z)DS \\ (x - \bar{x}_{nNN})Z(1 - D)(1 - S) \\ (x - \bar{x}_{nEE})Z(1 - D)S \\ (x - \bar{x}_{cEE})\frac{p_{01|0}-p_{01|1}}{p_{01|0}} - \bar{x}_{nEE}\frac{p_{01|1}}{p_{01|0}})(1 - Z)(1 - D)S \\ (x - \bar{x}_{cNN})\frac{p_{10|1}-p_{10|0}}{p_{10|1}} - \bar{x}_{aNN}\frac{p_{10|0}}{p_{10|1}})ZD(1 - S) \\ (x - \bar{x}_{cNE})\frac{q_{1|1}-q_{1|0}}{p_{00|0}} - \bar{x}_{cNN}\frac{p_{10|1}-p_{10|0}}{p_{00|0}} - \bar{x}_{nNN}\frac{p_{00|1}}{p_{00|0}})(1 - Z)(1 - D)(1 - S) \\ (x - \bar{x}_{cNE})\frac{q_{1|1}-q_{1|0}}{p_{11|1}} - \bar{x}_{cEE}\frac{p_{01|0}-p_{01|1}}{p_{11|1}} - \bar{x}_{aEE}\frac{p_{11|0}}{p_{11|1}})ZDS \\ x - \sum_k \pi_k \bar{x}_k \end{bmatrix}$$

where $\{\bar{x}_k\} = \{\bar{x}_{aNN}, \bar{x}_{aEE}, \bar{x}_{nNN}, \bar{x}_{nEE}, \bar{x}_{cNN}, \bar{x}_{cEE}, \bar{x}_{cNE}\}$. By Law of the Iterated Expectation, $E[g(\{\bar{x}_k\})] = 0$ when evaluated at the true value of $\{\bar{x}_k\}$.

Alternatively, we could also write the moment functions for the proportions of all the strata and then estimate the model together with the average baseline characteristics simultaneously by GMM. However, such GMM estimators do not behave well in our data. Thus, in our application we first estimate the proportions of all the strata as described in the paper, and then we estimate all the average baseline characteristics given the estimated proportions. As seen in $g(\{\bar{x}_k\})$, for each variable, we have nine equations (eight derived from the conditional expectations defined by $\{Z, D, S\}$ plus one from the expectation for the entire sample) to identify seven means, i.e., $\{\bar{x}_k\}$. Since the standard errors obtained from this GMM model do not take into account the fact that the proportions of the strata are also estimated, we employ a 500-repetition bootstrap to calculate the standard errors of the estimated average baseline characteristics.

Table A1: Average Baseline Characteristics for the *cEE* and *cNE* Strata

	Entire Sample			Non-Hispanics		
	<i>cEE</i>	<i>cNE</i>	Difference	<i>cEE</i>	<i>cNE</i>	Difference
Female	.396** (.015)	.630** (.165)	-.234 (.174)	.390** (.016)	.544** (.134)	-.154 (.145)
Age at Baseline	18.44** (.056)	19.19** (.699)	-.749 (.735)	18.39** (.068)	19.18** (.592)	-.786 (.632)
White, Non-hispanic	.299** (.012)	.260** (.126)	.039 (.133)	.369** (.015)	.259** (.126)	.110 (.135)
Black, Non-Hispanic	.445** (.013)	.622** (.158)	-.177 (.166)	.550** (.015)	.624** (.136)	-.074 (.147)
Has Child	.161** (.011)	.229** (.112)	-.068 (.119)	.151** (.012)	.210* (.111)	-.059 (.119)
Number of children	.215** (.018)	.356* (.187)	-.141 (.200)	.209** (.019)	.280 (.179)	-.071 (.192)
Personal Education	10.22** (.040)	10.34** (.504)	-.123 (.529)	10.24** (.048)	10.27** (.402)	-.036 (.434)
Ever Arrested	.230** (.012)	.223* (.128)	.007 (.136)	.228** (.013)	.292** (.112)	-.064 (.121)
At Baseline						
Have job	.241** (.011)	.174 (.108)	.068 (.115)	.244** (.012)	.159 (.102)	.084 (.110)
Weekly hours worked	24.07** (.583)	25.27** (6.365)	-1.196 (6.766)	24.05** (.613)	25.23** (5.768)	-1.187 (6.160)
Weekly earnings	113.86** (3.987)	120.08** (39.90)	-6.219 (40.90)	115.48** (3.500)	142.57** (34.31)	-27.09 (36.51)
Had job, Prev. Yr.	.714** (.013)	.585** (.141)	.129 (.151)	.718** (.014)	.588** (.126)	.130 (.136)
Months employed,Prev.Yr.	4.346** (.122)	3.286** (1.201)	1.060 (1.280)	4.435** (.137)	2.935** (1.105)	1.500 (1.201)
Earnings, Prev.Yr.	3396.2** (128.63)	3136.2** (1185.0)	260.02 (1250.6)	3377.6** (128.55)	2879.7** (1009.5)	497.88 (1095.9)

Note: Numbers in parentheses are standard errors. ** and * denote estimate is statistically different from 0 at 5% and 10% level, respectively. Computations use design weights. Missing values for each of the baseline variables were imputed with the mean of the variable. Standard errors are calculated by a 500-repetition bootstrap.

5 Bounds under Different Definitions of Enrollment

Table A2 compares the bounds on the effects of $\ln(wage)$ at week 208 for the *cEE* stratum using different definitions of enrollment. The first row shows the specific time spans used in the related literature to define enrollment. We start with our entire sample with continuously non-missing labor market outcomes. The first column replicates the results from the entire sample in Table 5 in the paper. The second column shows the results obtained by using the embargo period to define enrollment, as in Schochet et al. (2001) and Burghardt et al. (2001). That is, $D_i = 1$ if an individual enrolled in JC within 156 weeks after randomization, and $D_i = 0$ otherwise. The third column presents the results in the setting of Frumento et al. (2012), where

enrollment is considered within the first six months after randomization, and control members who enrolled within that period are dropped. That is, $D_i = 1$ if an individual enrolled in JC within 26 weeks, and $D_i = 0$ otherwise; then, individuals with $D_i = 1$ and $Z_i = 0$ are dropped from the sample.

By comparing columns (1) and (2), the proportion of compliers differs by 3 percentage points, which equals approximately the proportion of late crossovers among control members. Consistent with the definition of enrollment in the NJCS reports, the proportions of cEE and cNN members are larger than in our case, and the proportions of aEE and aNN members are lower. The remaining aEE members in (2) (i.e., early crossovers) experience a higher average wage than those in (1) (2.070 v.s. 2.033), which suggests that late crossovers probably suffer from the lock-in effects of JC and thus have lower wages. The narrower bounds in (2) than in (1) are probably due to a larger trimming proportion in (2). Note that both bounds for $E[Y^*(1)|cEE]$ in Proposition 1 are constrained solutions in (1), while the lower bound is an unconstrained solution in (2) because the large estimated value of $E[Y^*(1)|aEE]$ (2.070) provides additional information only for the upper bound. The identified region for the effect in Proposition 1 also shrinks in (2), though it still covers a zero effect. Similar to (1), the bounds for the effect in Proposition 2 identify a positive effect, with the lower bound equal to .053.

Because control members who enrolled within the first six months after randomization are dropped from the sample, always takers are absent in the setting of Frumento et al. (2012). As illustrated in Table A3, the bounds for $E[Y^*(1)|cEE]$ in their setting are derived from a cell containing two strata. Thus, we use the trimming procedure in Zhang, Rubin and Mealli (2008) (hereafter ZRM) and Lee (2009) to partially identify the wage effect for the cEE stratum in column (3). By comparing columns (1) and (3), the proportion of compliers differs by 1 percentage point, and the ITT effect on employment stays almost the same. The proportions of nEE , nNN and cNN members are larger in (3) than in (1), but the proportion of cEE members is approximately the same. The estimated values of $E[Y^*(0)|nEE]$ and $E[Y^*(0)|cEE]$ remain similar in spite of the new composition of the nEE and cEE strata. By the trimming procedure in ZRM and Lee (2009), the identified regions for the effect in both propositions

move slightly toward the positive region.

In sum, the results in Table A2 suggest that our estimated bounds are robust to the three different definitions of enrollment.

Table A2: Bounds on $\ln(wage)$ for cEE Using Different Definitions of Enrollment

Enrollment Definitions	208 Weeks (1)	156 Weeks (2)	26 Weeks (3)
References	Table 5: Entire Sample	NJCS Reports	Frumento et al.
$E[D Z = 1] - E[D Z = 0]$	.694** (.007)	.724** (.006)	.708** (.006)
$E[S Z = 1] - E[S Z = 0]$	.041** (.011)	.041** (.011)	.042** (.011)
π_{aEE}	.016** (.002)	.006** (.001)	-
π_{nEE}	.158** (.005)	.159** (.005)	.176** (.005)
π_{cEE}	.391** (.010)	.400** (.010)	.389** (.010)
π_{cNE}	.041** (.011)	.041** (.011)	.042** (.011)
π_{aNN}	.028** (.003)	.006** (.001)	-
π_{nNN}	.104** (.004)	.105** (.004)	.116** (.004)
π_{cNN}	.261** (.007)	.282** (.006)	.276** (.006)
α_{cEE}	.872** (.023)	.894** (.024)	.903** (.025)
$E[Y^*(1) aEE]$	2.033** (.059)	2.070** (.068)	-
$E[Y^*(0) nEE]$	2.033** (.016)	2.032** (.016)	2.034** (.014)
$E[Y^*(0) cEE]$	1.972** (.015)	1.973** (.015)	1.970** (.015)
$\bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$	2.429** (.066)	2.464** (.077)	-
$\bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$	1.676** (.034)	1.658** (.039)	-
<i>Bounds under Monotonicity (Proposition 1):</i>			
$[LY_{1,cEE}, UY_{1,cEE}]$	[1.951, 2.102]	[1.954, 2.099]	[1.953, 2.101]
$[L_{cEE}, U_{cEE}]$	[-.022, .130]	[-.020, .126]	[-.018, .131]
<i>Bounds under Monotonicity and Mean Dominance (Proposition 2):</i>			
$[\bar{LY}_{1,cEE}, \bar{UY}_{1,cEE}]$	[2.027, 2.102]	[2.027, 2.099]	[2.027, 2.101]
$[\bar{L}_{cEE}, \bar{U}_{cEE}]$	[.055, .130]	[.053, .126]	[.057, .131]
Number of Observations	9090	9090	9032

Note: Outcome is measured at week 208 after randomization. Numbers in parentheses are standard errors. ** denotes estimate is statistically different from 0 at 5% level. Computations use weights. Standard errors are calculated by a 5,000-repetition bootstrap. The definitions of α_{cEE} and α_{cNE} are provided in Proposition 1.

Table A3: Principal Strata by Enrollment Definition in Frumento et al.

$Z = 0$				$Z = 1$			
D				D			
0				0			
1				1			
S	0	cNE, cNN, nNN	-	S	0	nNN	cNN
	1	cEE, nEE	-		1	nEE	cNE, cEE

6 Bias for $E[Y^*(0)|cEE]$ in the Presence of Defiers

We derive the bias for the identified term $E[Y^*(0)|cEE]$ when defiers exist. We find that the bias has a similar expression as the one for *LATE* in equation (18) of Angrist, Imbens and Rubin (1996).

Table A4 shows the distribution of strata in the observed data when only Assumption 4 fails. By equation (4) in the paper, the identified term $E[Y^*(0)|cEE]$ equals:

$$E[\widehat{Y^*(0)}|cEE] = \frac{p_{01|0}}{p_{01|0} - p_{01|1}} \bar{Y}^{001} - \frac{p_{01|1}}{p_{01|0} - p_{01|1}} \bar{Y}^{101}.$$

Without Assumption 4, this quantity becomes:

$$\begin{aligned} E[\widehat{Y^*(0)}|cEE] &= \frac{1}{\pi_{cEE} - \pi_{dEE}} \{ (\pi_{cEE} E[Y^*(0)|cEE] + \pi_{nEE} E[Y^*(0)|nEE]) \\ &\quad - (\pi_{dEE} E[Y^*(0)|dEE] + \pi_{nEE} E[Y^*(0)|nEE]) \} \\ &= E[Y^*(0)|cEE] + \underbrace{\frac{\pi_{dEE}}{\pi_{cEE} - \pi_{dEE}} (E[Y^*(0)|cEE] - E[Y^*(0)|dEE])}_{Bias} \end{aligned}$$

Thus, the bias for the identified term $E[Y^*(0)|cEE]$ also depends on the ratio $\frac{\pi_{dEE}}{\pi_{cEE} - \pi_{dEE}}$ and the difference in the average outcomes between compliers and defiers.

Table A4: Distribution of Principal Strata in the Presence of Defiers

$Z = 0$				$Z = 1$			
D				D			
		0	1			0	1
S	0	cNE, cNN, nNN	aNN, dNN	S	0	dNN, dEN, nNN	cNN, aNN
	1	cEE, nEE	aEE, dEE, dEN		1	dEE, nEE	cNE, cEE, aEE

7 Additional Tables for the JC Application

Table A5: Summary Statistics of Baseline Variables

	Entire Sample				Non-Hispanics			
	Missing Prop.	Z=1	Z=0	Diff.(Std.Err.)	Missing Prop.	Z=1	Z=0	Diff.(Std.Err.)
Female	0	.454	.458	-.004 (.011)	0	.454	.454	-.010 (.012)
Age at Baseline	0	18.44	18.35	.087* (.046)	0	18.44	18.34	.096* (.050)
White, Non-hispanic	0	.265	.263	.002 (.009)	0	.319	.318	.001 (.011)
Black, Non-Hispanic	0	.494	.491	.003 (.011)	0	.595	.593	.002 (.012)
Hispanic	0	.169	.172	-.003 (.008)	—	—	—	—
Other Race/Ethnicity	0	.072	.074	-.002 (.006)	0	.087	.089	-.003 (.007)
Never married	.017	.917	.916	.002 (.006)	.018	.926	.924	.002 (.006)
married	.017	.020	.023	-.003 (.003)	.018	.015	.018	-.003 (.003)
Living together	.017	.039	.040	-.002 (.004)	.018	.035	.037	-.002 (.004)
Separated	.017	.024	.021	.003 (.003)	.018	.023	.020	.003 (.003)
Has Children	.007	.189	.193	-.004 (.008)	.006	.187	.190	-.004 (.009)
Number of children	.010	.270	.268	.002 (.014)	.180	.269	.271	-.003 (.015)
Education (years of schooling)	.018	10.12	10.11	.013 (.033)	.018	10.14	10.12	.022 (.036)
Mother's Education	.188	11.49	11.46	.030 (.061)	.182	11.81	11.83	-.021 (.055)
Father's Education	.383	11.40	11.54	-.145* (.077)	.379	11.72	11.86	-.147** (.072)
Ever Arrested	.017	.248	.249	-.001 (.009)	.017	.255	.257	-.002 (.010)
Household Income: <3000	.358	.253	.251	.001 (.012)	.357	.248	.244	.004 (.013)
3000-6000	.358	.205	.208	-.003 (.011)	.357	.202	.213	-.012 (.012)
6000-9000	.358	.117	.114	.003 (.009)	.357	.119	.105	.015 (.009)
9000-18000	.358	.246	.245	.001 (.011)	.357	.244	.248	-.003 (.013)
>18000	.358	.180	.182	-.002 (.010)	.357	.187	.191	-.003 (.012)
Personal Income: <3000	.079	.788	.789	-.001 (.009)	.077	.787	.788	-.001 (.010)
3000-6000	.079	.128	.131	-.003 (.008)	.077	.129	.136	-.007 (.008)
6000-9000	.079	.053	.046	.007 (.005)	.077	.052	.043	.009* (.005)
>9000	.079	.031	.034	-.003 (.004)	.077	.031	.033	-.001 (.004)
At Baseline:								
Have job	.021	.198	.192	.007 (.009)	.021	.204	.187	.017* (.009)
Weekly hours worked	0	21.83	20.91	.922** (.447)	0	21.97	20.75	1.216** (.491)
Weekly earnings	0	111.08	102.89	8.183 (5.134)	0	107.79	102.28	5.516** (2.804)
Had job, Prev. Yr.	.017	.635	.627	.008 (.010)	.017	.642	.627	.015 (.011)
Months employed,Prev.Yr.	0	3.603	3.530	.074 (.091)	0	3.654	3.512	.143 (.100)
Earnings, Prev.Yr.	.062	2911.0	2810.5	100.56 (117.58)	.064	2900.3	2794.7	105.57 (106.34)
Numbers of observations	9090	5491	3599		7529	4551	2978	

Note: Z is an indicator for whether the individual was randomly assigned to participate or not in JC. ** and * denote difference is statistically different from 0 at 5% and 10% level, respectively. Computations use design weights.

Table A6: Bounds for the Effect of JC for cEE Adjusting for Non-Response

	Entire Sample	Non-Hispanics
π_{aEE}	.017** (.002)	.018** (.002)
π_{nEE}	.162** (.005)	.162** (.005)
π_{cEE}	.391** (.009)	.381** (.010)
π_{cNE}	.038** (.010)	.049** (.011)
π_{aNN}	.026** (.003)	.028** (.003)
π_{nNN}	.108** (.004)	.108** (.004)
π_{cNN}	.258** (.006)	.254** (.007)
α_{cEE}	.877** (.022)	.851** (.024)
$E[Y^*(1) aEE]$	2.010** (.050)	2.001** (.052)
$E[Y^*(0) nEE]$	2.033** (.015)	2.032** (.016)
$E[Y^*(0) cEE]$	1.985** (.013)	1.961** (.015)
$\bar{Y}(y_{\alpha_{cEE}}^{111} \leq Y \leq y_{1-\alpha_{cNE}}^{111})$	2.449** (.065)	2.383** (.056)
$\bar{Y}(y_{\alpha_{cNE}}^{111} \leq Y \leq y_{1-\alpha_{cEE}}^{111})$	1.666** (.036)	1.702** (.034)
<i>Bounds under Monotonicity (Proposition 1):</i>		
$[LY_{1,cEE}, UY_{1,cEE}]$	[1.957, 2.102]	[1.940, 2.114]
CLR 95% confidence interval	(1.928, 2.127)	(1.911, 2.140)
$[L_{cEE}, U_{cEE}]$	[-.028, .117]	[-.021, .153]
CLR 95% confidence interval	(-.065, .153)	(-.060, .191)
<i>Bounds under Monotonicity and Mean Dominance (Proposition 2):</i>		
$[\bar{LY}_{1,cEE}, \bar{UY}_{1,cEE}]$	[2.029, 2.102]	[2.028, 2.114]
CLR 95% confidence interval	(2.013, 2.128)	(2.011, 2.142)
$[\bar{L}_{cEE}, \bar{U}_{cEE}]$	[.044, .117]	[.067, .153]
CLR 95% confidence interval	(.015, .155)	[.035, .193]
Number of observations	10520	8701

Note: Outcome is measured at week 208 after randomization. Numbers in parentheses are standard errors. ** denotes estimate is statistically different from 0 at 5% level. Computations use weights accounting for sample design, interview design and interview non-response. Standard errors are calculated by a 5,000-repetition bootstrap. Numbers in square brackets are half-median unbiased estimates of the bounds, and the numbers below them are CLR 95% confidence intervals, which contain the true value of the parameter with a given probability (see Section 2.5 for details). The definitions of α_{cEE} and α_{cNE} are provided in Proposition 1.

8 Additional Table and Figures for the Simulation Study

Table A7: Simulation Summary

Mimic characteristics of the Entire Sample in Table 5:		
General	9090 obs., each randomly assigned to $Z = 1$ w.p. 5491/9090; membership of principal strata: $\mathcal{U}(0,1)$	
simulation	Observed D and S are jointly determined by membership of principal strata and Z	
design	Wages: $\ln \mathcal{N}(\mu_{d,k}, \sigma^2)$: $\mu_{1,cEE}=2.04$, $\mu_{0,cEE}=1.97$, $\mu_{1,aEE}=2.035$, $\mu_{0,aEE}=2.035$, $\mu_{1,dEE}=2.02$, $\mu_{0,dEE}=1.99$, $\mu_{1,dEN}=1.85$, $\sigma^2=.2$; True Effect (Δ) =.07; values are chosen s.t. they are close to observed $\bar{Y}^{zds}$	
Specific design	<i>Version 1</i> (A3 and A4 hold: $\pi_c=.695$, $\pi_d=0$)	
	<i>Version 2</i> (A5 and A6 hold: $\pi_{cEN}=0$, $\pi_{dNE}=0$, $\mu_{1,cNE}=1.88$)	
Mimic observed $p_{ds z}$:	Mimic ratio a/n and composition within a , n , c and d :	
	$\pi_{aNN}=.03$, $\pi_{aEE}=.015$, $\pi_{nNN}=.105$, $\pi_{nEE}=.155$ $\pi_{cEE}=.395$, π_{cEN} , $\pi_{cNN}=.26$, π_{cEN} , $\pi_{cNE}=.04$, π_{cEN} $\frac{\pi_n}{\pi_a}=5.7$, $\frac{\pi_{aEE}}{\pi_a}=.3$, $\frac{\pi_{nEE}}{\pi_n}=6$, $\frac{\pi_{cEE}}{\pi_c}=.559$, $\frac{\pi_{cNE}}{\pi_c}=.059$, $\frac{\pi_{dEE}}{\pi_d}=.5$, $\frac{\pi_{dEN}}{\pi_d}=1$; $E[D(1)-D(0)]=\pi_c-\pi_d$	
Variables	<i>Simulation 1</i> (A5 fails)	<i>Simulation 2</i> (A6 fails)
	<i>Simulation 3</i> (A5 and A6 fail)	<i>Simulation 4</i> (A3 fails)
Constants	<i>Simulation 5</i> (A3 and A4 fail)	
	π_{cEN} , thus π_{cEE} , π_{cNN} , π_{cNE} $\mu_{1,cNE}=1.88$	
General conclusion	$\text{LB1} < \Delta < \text{UB}$ as π_{cEN} $<.15$; $\text{LB2} > 0$ as π_{cEN} $<.12$; at $\pi_{cEN}=0$, close to empirical ones	
	$\text{LB1} < \Delta < \text{UB}$ over symmetric range near 0, $\text{LB2} < \Delta < \text{UB}$ over a shrinking range as π_{cEN} rises; $\text{LB2} > 0$ over a smaller range; Δ crosses LB2 at 0	
General conclusion	$\text{LB1} < \Delta < \text{UB}$ over [-5, .5]; Prob. that $\text{LB2} < \Delta < \text{UB}$ rises as $\mu_{1,cEE}-\mu_{1,cNE}$ increases; Δ crosses LB2 at 0	
	Huge variations with a weak IV; as IV > .4, p5th of UB is above p95th of LB1, bounds are valid w.p. >50%; at IV=.695, close to empirical ones	
General conclusion	Prob. that Δ falls within bounds collapses as IV gets weaker; IV=.695: $\text{LB2} < \Delta < \text{UB}$ as $\pi_d < .15$; IV=.4: Δ crosses LB2 at $\pi_d=.12$, p95th and p5th overlap across bounds; IV=.02: huge variations	
	$E[D(1)-D(0)]=.695$, .4, .02, and π_d , thus π_c , π_{cEE}	

Notes: UB denotes the average of the estimated upper bounds across 1000 replications in Proposition 1, and LB1 and LB2 denote the average of the estimated lower bounds in Proposition 1 and in Proposition 2, respectively. p5th and p95th denote the 5th and 95th percentiles of the estimates, respectively.

Figure 1: Results When Only Assumption 5 Fails (Simulation 1). The solid line plots the true effect (TE), the dashed lines plot the mean of the estimates of L_{cEE} and U_{cEE} across 1000 replications, and the dotted line plots that of $\bar{L}_{cEE}$. The 5th and 95th percentiles of the estimates of U_{cEE} are plotted as circles, and those for L_{cEE} and $\bar{L}_{cEE}$ are plotted as squares and triangles, respectively.

Figure 2: Results When Only Assumption 6 Fails (Simulation 2). See Figure 1 for a description of the plots.

Figure 3: Results When Assumptions 5 and 6 Fail (Simulation 3). See Figure 1 for a description of the plots.

Figure 4: Results Regarding the Strength of the IV (Simulation 4). See Figure 1 for a description of the plots.

Figure 5: Results When Assumption 4 Fails for Different IV Strengths (Simulation 5). See Figure 1 for a description of the plots.

References

- [1] Angrist, J., Imbens, G. and Rubin, D. (1996), "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444-472.
- [2] Burghardt, J., Schochet, P.Z., McConnell, S., Johnson, T., Gritz, R.M., Glazerman, S., Homrighausen, J. and Jackson, R. (2001), *Does Job Corps Work? Summary of the National Job Corps Study*, 8140-530, Mathematica Policy Research, Inc., Princeton, NJ.
- [3] Chernozhukov, V., Lee, S. and Rosen, A. (2013), "Intersection Bounds: Estimation and Inference," *Econometrica*, 81(2), 667-737.
- [4] Frumento, F., Mealli, F., Pacini, B. and Rubin, D. (2012), "Evaluating the Effect of Training on Wages in the Presence of Noncompliance, Nonemployment, and Missing Outcome Data," *Journal of the American Statistical Association*, 107 (498), 450-466.
- [5] Horowitz, J.L. and Manski, C.F. (1995), "Identification and Robustness with Contaminated and Corrupted Data," *Econometrica*, 63 (2), 281-302.
- [6] Imbens, G. and Manski, C.F. (2004), "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72 (6), 1845-1857.
- [7] Lee, D. (2009), "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071-102.
- [8] Schochet, P., Burghardt, J. and Glazerman, S. (2001), *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes*, 8140-530, Mathematica Policy Research, Inc., Princeton, NJ.
- [9] Zhang, J.L., Rubin, D. and Mealli, F. (2008), "Evaluation the Effects of Job Training Programs on Wages Through Principal Stratification," in D. Millimet et al. (eds) *Advances in Econometrics*, XXI, Elsevier.