SUPPLEMENT TO "HOROWITZ-MANSKI-LEE BOUNDS WITH MULTILAYERED SAMPLE SELECTION"

KORY KROFT*, ISMAEL MOURIFIÉ †, AND ATOM VAYALINKAL‡

Appendix A. Additional Results

A.1. **Bounds on Aggregate Effect for "Stayers".** To derive bounds on the aggregate LCDE, defined as $\sum_{d=l}^{l'} \frac{p_{d,d}}{\sum_{d'=l}^{l'} p_{d',d'}} \mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$, one might be tempted to adopt a naive approach by taking a weighted average of the pointwise sharp bounds derived in Theorem 1(i). However, this approach not only fails to provide sharp bounds on the aggregate quantity but may also yield invalid bounds. The primary issue lies in the fact that both the weights and the LCDE quantities (i.e. the $\mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$ terms) are only set-identified. Consequently, the naive aggregation does not preserve the sharpness or validity of the bounds. To better illustrate this problem, consider the following.

The sharp bounds on the aggregate LCDE can be tighter than the weighted average of the marginal bounds on $\mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$ for $d = l, \ldots, l'$ when these marginal bounds are achieved at values of $p_{d,d}$ for $d = l, \ldots, l'$ that are not *jointly* attainable. Specifically, this occurs whenever $\left(\underline{p}_{d,d} : d = l, \ldots, l'\right)$ does not lie within the identified set for $(p_{d,d} : d = l, \ldots, l')$. As a result, the sharp bounds, which are obtained by optimizing over the identified set for the *vector* $(p_{d,d} : d = l, \ldots, l')$, may be tighter.

The sharp bounds on the LCDE can be wider than the weighted average of the marginal LCDE bounds when there exist $d, d' \in \{l, \ldots, l'\}$ such that the lower bound for $\mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$ is significantly smaller than the lower bound for $\mathbb{E}[Y_{1,d'} - Y_{0,d'} \mid T = (d',d')]$, and $\underline{p}_{d,d}$ is also much smaller than $\underline{p}_{d',d'}$.

In such cases, the choice of $p_{d,d}$ directly affects the relative weight of $\mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$ in the aggregate LCDE. Consequently, the value of the objective function in the optimization problem defining the lower bound in Theorem 1(iii) may actually decrease as $p_{d,d}$ increases.

To illustrate this more clearly, consider a simple example with only two groups, a and b. For each $d \in \{a,b\}$, let $L_d(p)$ denote the sharp lower bound for $\mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$ under the assumption that $p_{d,d} = p$. Suppose $\underline{p}_{b,b} > \underline{p}_{a,a} = 0$, so that the lower bound for $\mathbb{E}[Y_{1,a} - Y_{0,a} \mid T = (a,a)]$, given by $L_a(\underline{p}_{a,a}) = L_a(0) = y_L$, is the trivial bound. Additionally, assume the lower bound for $\mathbb{E}[Y_{1,b} - Y_{0,b} \mid T = (b,b)]$ is non-trivial, i.e., $L_b(\underline{p}_{b,b}) > y_L$.

In this scenario, the weighted average of these lower bounds simplifies to:

$$\frac{\underline{p}_{a,a} L_a(\underline{p}_{a,a}) + \underline{p}_{b,b} L_b(\underline{p}_{b,b})}{\underline{p}_{a,a} + \underline{p}_{b,b}} = L_b(\underline{p}_{b,b}),$$

since $\underline{p}_{a,a} = 0$.

Now, consider the case where there exists a point $p > 0$ such that $(p, \underline{p}_{b,b})$ lies within the identified set for $(p_{a,a}, p_{b,b})$, and $L_a(p) < L_b(\underline{p}_{b,b})$. In this case, we have:

$$\frac{p L_a(p) + \underline{p}_{b,b} L_b(\underline{p}_{b,b})}{p + \underline{p}_{b,b}} < L_b(\underline{p}_{b,b}).$$

By definition, the sharp lower bound given in Theorem 1(iii) will be at least as small as the left-hand side above, resulting in a smaller lower bound for the aggregate LCDE. Similarly, it is also possible for the sharp upper bound to be larger than the weighted average of the marginal upper bounds.

Appendix B. Relation to the Mediation Analysis Literature

B.1. **Direct and Indirect Effects in Presence of Sample Selection.** This section establishes a connection between our model and the literature on mediation analysis (Pearl, 2001). In mediation analysis, where Z represents the randomized treatment, D is the "mediator", and Y is the outcome, the treatment (job training) can influence the outcome through two distinct channels: a direct channel and an indirect channel that passes through the mediator. The vector W of latent unobserved variables (often called "confounding variables") simultaneously affect D and Y, making D an endogenous variable. In our setting, the mediator corresponds to the firm where the individual would be employed if they were externally assigned to job training. The graphical representation of the outcome equation in our model takes the form:[1]

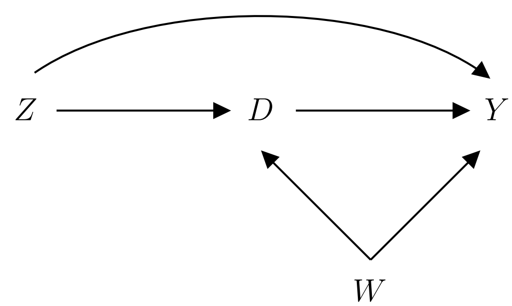


FIGURE 1. DAG of causal relationships between variables in our model.

The mediation literature distinguishes two causal estimands: direct and indirect effects. The *control direct effect* (CDE) is defined as:

$$\text{CDE}(d) \equiv \mathbb{E}[Y_{1,d} - Y_{0,d}]. \tag{B.1}$$

The CDE captures the causal effect of job training on earnings when firm type d is held fixed —equivalently, the within-firm wage effect for layer d. In Lee's (2009) terminology, the CDE corresponds to the causal impact of job training on the wage rate illustrated by the curved arrow in Figure 1. This parameter is the primary focus of Lee (2009).

CDEs are useful when the policymaker is primarily interested in the impact of job training on wages at a specific firm. Policymakers may also be interested in understanding the overall impact of training on wages at firms that workers naturally choose when they receive training. The second type of direct effect – the "natural direct effect" (NDE) – captures this notion:

$$\text{NDE} \equiv \mathbb{E}[Y_{1,D_1} - Y_{0,D_1}] = \sum_{d=0}^{K} \mathbb{E}[Y_{1,d} - Y_{0,d} | D_1 = d] \times \mathbb{P}[D_1 = d] \tag{B.2}$$

[1]For the sake of clarity, this graph simplifies the discussion by omitting sample selection.

where $Y_{z,D_{z'}} \equiv \sum_{d=0}^{K} Y_{z,d} 1\{D_{z'} = d\}$ for $z, z' \in \{0,1\}$, and $d \in \{0, ..., K\}$. The expression $Y_{1,D_1} - Y_{0,D_1}$ is the wage impact of job training at the firm a worker would select under treatment assignment. The NDE averages these individual effects.

Turning to indirect effects, the "natural indirect effect" (NIE) is defined as:

$$\begin{aligned} \text{NIE} &\equiv \mathbb{E}[Y_{0,D_1} - Y_{0,D_0}] \\ &= \sum_{d=0}^{K} \sum_{d'=0:d'\neq d}^{K} \mathbb{E}[Y_{0,d} - Y_{0,d'} | D_0 = d', D_1 = d] \times \mathbb{P}[D_0 = d', D_1 = d]. \quad \text{(B.3)} \end{aligned}$$

The term $Y_{0,d} - Y_{0,d'}$ represents the wage gap between firms d and d' absent job training. Evaluated at the natural representative firms D_1 and D_0, this becomes $Y_{0,D_1} - Y_{0,D_0}$. The indirect effect isolates the impact of job training operating purely through firm change. As highlighted by Pearl (2009), this estimand is empirically controversial since suppressing the direct effect of Z on Y while preserving the indirect channel is not realistic, though it remains standard in the mediation literature.

We now introduce two key parameters, the Local Controlled Direct Effect (LCDE) and the Local Controlled Indirect Effect (LCIE):

$$\text{LCDE}(d|t) = \mathbb{E}[Y_{1,d} - Y_{0,d} | T = t], d \in \{1, ..., K\}, \text{ and } t \in \mathcal{T} \quad \text{(B.4)}$$

$$\text{LCIE}(z, d, d'|t) = \mathbb{E}[Y_{z,d} - Y_{z,d'} | T = t], d \in \{1, ..., K\}, \text{ and } t \in \mathcal{T} \quad \text{(B.5)}$$

The LCDE allows the CDE to vary across response types t, accommodating individual-level heterogeneity in firm-specific treatment effects. The LCDE can be more policy-relevant than the CDE in specific contexts, mirroring the ATE vs. LATE debate in the IV literature. The same applies to the LCIE.

The "sample selection" analogues of the CDE, NDE, and NIE defined conditionally on always-employed workers $\{D_0 > 0, D_1 > 0\}$ are weighted average of LCDE$(d|t)$ or

LCIE$(z, d, d'|t)$.

$$\mathbb{E}[Y_{1,d} - Y_{0,d}|D_0 > 0, D_1 > 0] = \sum_{l=1}^{K}\sum_{l'=1}^{K} \text{LCDE}(d|l,l') \times \mathbb{P}[T = (l,l')|D_0 > 0, D_1 > 0],$$

$$\mathbb{E}[Y_{1,D_1} - Y_{0,D_1}|D_0 > 0, D_1 > 0] = \sum_{d=1}^{K}\sum_{d'=1}^{K} \text{LCDE}(d|d',d) \times \mathbb{P}[T = (d',d)|D_0 > 0, D_1 > 0],$$

$$\mathbb{E}[Y_{0,D_1} - Y_{0,D_0}|D_0 > 0, D_1 > 0] = \sum_{d=1}^{K}\sum_{d'=1:d\neq d'}^{K} \text{LCIE}(0,d,d'|d',d) \times \mathbb{P}[T = (d',d)|D_0 > 0, D_1 > 0].$$

Appendix C. 2 Firm Types Case: A numerical illustration

This section considers our model two firm types: high type (H) and low type (L). Under Assumption 2, the always-employed (AE) consist of four response types:

$$\{D_0 > 0, D_1 > 0\} = \{(L,L), (H,H), (L,H), (H,L)\} \equiv AE.$$

To establish bounds on our causal effects of interest, we first characterize identification of the response-type probabilities: $\{p_t : t \in \mathcal{T}\}$. Using information on (D, Z) only, $\{p_t : t \in \mathcal{T}\}$ has to satisfy equations (4.4, 4.5) from step 1 above. In addition, if we impose Assumption 2, i.e., $\mathcal{R}_T = \{\text{Assumption 2}\}$, the identified set for the response types in this simple case is characterized by non-negative solutions to the following set of (in)equalities:

$$p_{0,0} = 1 - \mathbb{P}(D = H \mid Z = 1) - \mathbb{P}(D = L \mid Z = 1), \tag{C.1}$$

$$p_{0,L} = \mathbb{P}(D = L \mid Z = 1) - \mathbb{P}(D = H \mid Z = 0) + p_{H,H} - p_{L,L}, \tag{C.2}$$

$$p_{0,H} = \mathbb{P}(D = H \mid Z = 1) - \mathbb{P}(D = L \mid Z = 0) + p_{L,L} - p_{H,H}, \tag{C.3}$$

$$p_{L,H} = \mathbb{P}(D = L \mid Z = 0) - p_{L,L}, \tag{C.4}$$

$$p_{H,L} = P(D = H \mid Z = 0) - p_{H,H}, \tag{C.5}$$

$$\max\{0, \mathbb{P}(D = H|Z = 0) - \mathbb{P}(D = L|Z = 1)\} \leq p_{H,H} \leq \min\{\mathbb{P}(D = H|Z = 0), \mathbb{P}(D = H|Z = 1)\} \tag{C.6}$$

$$\max\{0, \mathbb{P}(D = L|Z = 0) - \mathbb{P}(D = H|Z = 1)\} \leq p_{L,L} \leq \min\{\mathbb{P}(D = L|Z = 0), \mathbb{P}(D = L|Z = 1)\}. \tag{C.7}$$

More precisely, we can show that

$$\Theta_I(\mathcal{R}_T) = \{(p_t : t \in \mathcal{T}) \geq 0 : \text{ such that eqs } (C.1) \text{ to } (C.7) \text{ are satisfied}\}.$$

In this case, it is also straightforward to show that $\Theta_I(\mathcal{R}_T) \neq \emptyset$ if and only if the propensity scores satisfy

$$\mathbb{P}(D = 0|Z = 1) \leq \mathbb{P}(D = 0|Z = 0). \tag{C.8}$$

In this particular case, we have:

$$\underline{\gamma}^z_{H,H} = \frac{\max\{0, \mathbb{P}(D = H|Z = 0) - \mathbb{P}(D = L|Z = 1)\}}{\mathbb{P}(D = H|Z = z)}, \text{ for } z \in \{0,1\},$$
$$\underline{\gamma}^z_{L,L} = \frac{\max\{0, \mathbb{P}(D = L|Z = 0) - \mathbb{P}(D = H|Z = 1)\}}{\mathbb{P}(D = L|Z = z)}, \text{ for } z \in \{0,1\},$$

In the numerical illustration below, we describe $\Theta_I(\mathcal{R}_T)$ and demonstrate how imposing further assumptions on response types can significantly refine $\Theta_I(\mathcal{R}_T)$. In this case, as shown by (C.1)-(C.5), the response-type probability p_t, for each $t \in \{(0,0),(0,L),(0,H),(L,H),(H,L)\}$, can be represented as a linear function of $p_{H,H}$ and $p_{L,L}$, which are themselves only set identified; in other words, $\Theta_I(\mathcal{R}_T)$ is parameterized by $(p_{H,H}, p_{L,L})$. Hence, our forthcoming discussion will focus primarily on illustrating the projection of $\Theta_I(\mathcal{R}_T)$ with respect to the coordinates of $(p_{H,H}, p_{L,L})$.

C.0.1. *Data Generating Process.* We first generate propensity scores to be consistent with the data in Lee (2009). This is reported in Table 1.

TABLE 1. Propensity scores for numerical illustrations

Job Training	$P(D = L\|Z = z)$	$P(D = H\|Z = z)$
$Z = 1$	0.302886	0.408114
$Z = 0$	0.313959	0.373041

The data are generated such that the true values for $p_{H,H}$, and $p_{L,L}$ are:

$$p_{H,H} = \mathbb{P}(D = H \mid Z = 0) = 0.373041,$$
$$p_{L,L} = 0.278886.$$

Next, we randomly generate the outcomes.[2] For each type $t \in \mathcal{T}$, let D_z denote employment status when externally assigned $Z = z$. The conditional distributions of $\exp(Y_{z,D_z}) \mid T = t$ for each $t \in \mathcal{T}$ are assumed to follow Lognormal$(\mu_{z|t}, \sigma_{z|t})$ distributions. Here, $\sigma_{z|t} = 1$ for all combinations of z and t, indicating that variability only arises through $\mu_{z|t}$ between types. We present two distinct potential earnings distribution models as described in Table 2.

Table 2. Potential Outcome distributions for the simulated DGPs

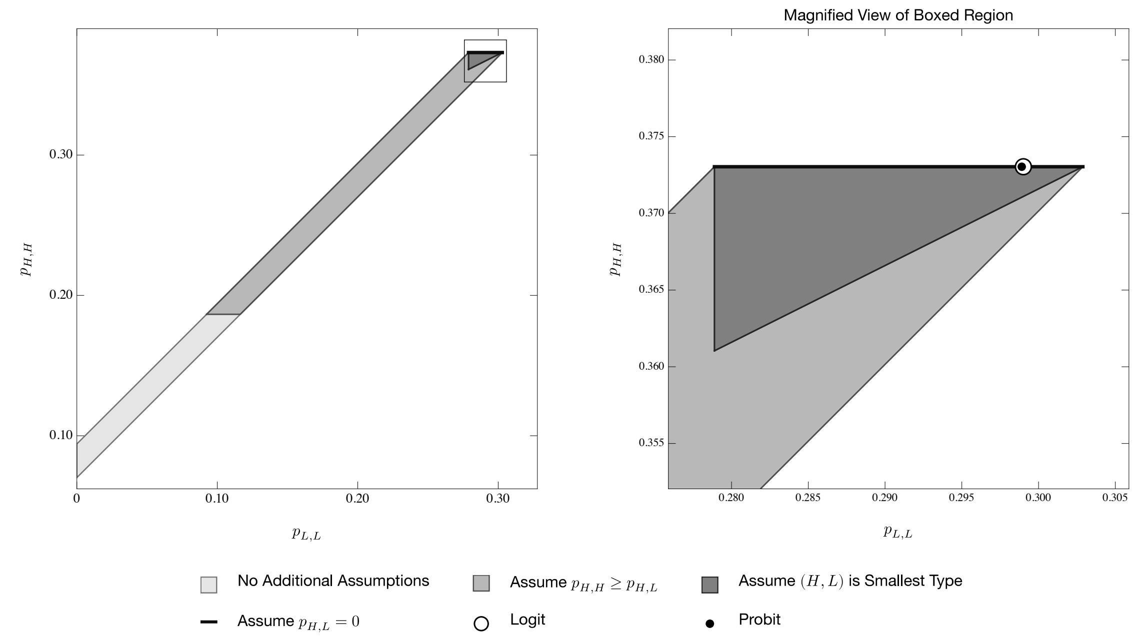
	Design 1		Design 2	
t	$\exp(Y_{1,D_1}) \mid T = t$	$\exp(Y_{0,D_0}) \mid T = t$	$\exp(Y_{1,D_1}) \mid T = t$	$\exp(Y_{0,D_0}) \mid T = t$
$(0, L)$	Lognormal$(9.5, 1)$	Lognormal$(9.5, 1)$	Lognormal$(10.5, 1)$	Lognormal$(9.5, 1)$
$(0, H)$	Lognormal$(11.5, 1)$	Lognormal$(9.5, 1)$	Lognormal$(12.5, 1)$	Lognormal$(9.5, 1)$
(L, H)	Lognormal$(16.5, 1)$	Lognormal$(9.5, 1)$	Lognormal$(14.5, 1)$	Lognormal$(9.5, 1)$
(H, L)	Lognormal$(9.75, 1)$	Lognormal$(9.6, 1)$	Lognormal$(10.5, 1)$	Lognormal$(10.5, 1)$
(L, L)	Lognormal$(9.5, 1)$	Lognormal$(9.5, 1)$	Lognormal$(10.5, 1)$	Lognormal$(9.5, 1)$
(H, H)	Lognormal$(14.5, 1)$	Lognormal$(14.5, 1)$	Lognormal$(14, 1)$	Lognormal$(12, 1)$

C.0.2. *Simulations Results.* We begin by exploring the geometry of $\Theta_I(\mathcal{R}_T)$, and demonstrate how incorporating further assumptions regarding response types can significantly shrink its shape.

Figure 3, along with Tables 3 and 4, present the results for both designs. Initially, we compute Lee bounds, which set identifies the total effect $\mathbb{E}(Y_{1,D_1} - Y_{0,D_0}|D_0 > 0, D_1 > 0)$, as in Lemma 2. Within the framework of design 1, these bounds lie entirely within the positive quadrant, excluding 0, suggesting that job training increases wage rates for the always-employed AE. However, for this DGP, the within-firm effect is 0. As illustrated below, whenever $\mathbb{E}(Y_{1,H} - Y_{0,H}|T = (H, H)) = \mathbb{E}(Y_{1,L} - Y_{0,L}|T = (L, L)) = 0$, Lee bounds primarily capture the sorting effect:

$$\mathbb{E}(Y_{1,D_1} - Y_{0,D_0}|D_0 > 0, D_1 > 0) = \frac{p_{L,H}}{\mathbb{P}(AE)}\mathbb{E}[Y_{1,H} - Y_{0,L}|T = (L, H)] + \frac{p_{H,L}}{\mathbb{P}(AE)}\mathbb{E}[Y_{1,L} - Y_{0,H}|T = (H, L)].$$

[2]By construction, the outcomes (wages) simulated here are independent of the dataset used in Lee (2009).



This figure plots the identified set for $(p_{L,L}, p_{H,H})$ under various assumptions, as indicated, for our simulation designs (the plots are identical for both designs). The right panel is a magnified view of the boxed region in the left panel. Lighter regions correspond to weaker restrictions and contain the darker regions. Under the Logit and Probit assumptions, $(p_{L,L}, p_{H,H})$ is point-identified.

FIGURE 2. Identified set for $(p_{L,L}, p_{H,H})$ in both simulated DGPs.

TABLE 3. Multilayered for Simulation Designs.

			$\mathbb{E}(Y_{1,H} - Y_{0,H} \mid T = (H,H))$		$\mathbb{E}(Y_{1,L} - Y_{0,L} \mid T = (L,L))$	
Design 1	$p^*_{H,H}$	$p^*_{L,L}$	lower	upper	lower	upper
Baseline	0.0702	0.0000	-2.8749	3.4001	Trivial Bounds	
$p_{H,H} \geq p_{H,L}$	0.1865	0.0924	-1.5956	1.9588	-2.3438	2.2672
(H, L) is smallest type	0.3610	0.2789	-0.1698	0.4873	-0.3958	0.3458
$p_{H,L} = 0$	0.3730	0.2789	-0.0382	0.3591	-0.3958	0.3458
Logit	0.3730	0.2990	-0.0382	0.3591	-0.1628	0.1095
Probit	0.3730	0.2989	-0.0382	0.3591	-0.1643	0.1110
Design 2						
Baseline	0.0702	0.0000	-0.8954	4.9799	Trivial Bounds	
$p_{H,H} \geq p_{H,L}$	0.1865	0.0924	0.3699	3.7242	-1.3438	3.2672
(H, L) is smallest type	0.3610	0.2789	1.7503	2.3440	0.6042	1.3458
$p_{H,L} = 0$	0.3730	0.2789	1.8743	2.2196	0.6042	1.3458
Logit	0.3730	0.2990	1.8743	2.2196	0.8372	1.1095
Probit	0.3730	0.2989	1.8743	2.2196	0.8357	1.1110

Notes: Outcome is simulated log earnings. p^*_t is the minimum value of p_t over the identified set for response-types under the given assumption.

Table 4. Aggregate multilayered bounds for numerical illustrations.

	p^*_{HH}	p^*_{LL}	p^*_{HH}	p^*_{LL}	$\sum_{d\in\{L,H\}} \frac{p_{d,d}}{p_{HH}+p_{LL}} \mathbb{E}[Y_{1,d}-Y_{0,d}\mid T=(d,d)]$	
Design 1	(for lower bound)		(for upper bound)		lower	upper
Baseline	0.0758	0.0056	0.0723	0.0021	-2.9349	3.4206
$p_{H,H} \geq p_{H,L}$	0.1865	0.0924	0.1865	0.0924	-1.8434	2.0609
(H,L) is smallest type	0.3610	0.2789	0.3610	0.2789	-0.2683	0.4256
$p_{H,L}=0$	0.3730	0.2789	0.3730	0.2789	-0.1912	0.3534
Logit	0.3730	0.2990	0.3730	0.2990	-0.0937	0.2481
Probit	0.3730	0.2989	0.3730	0.2989	-0.0943	0.2487
Design 2						
Baseline	0.0823	0.0122	0.0715	0.0013	-1.0519	4.9918
$p_{H,H} \geq p_{H,L}$	0.1865	0.0924	0.1865	0.0924	-0.1977	3.5728
(H,L) is smallest type	0.3610	0.2789	0.3610	0.2789	1.2509	1.9090
$p_{H,L}=0$	0.3730	0.2789	0.3730	0.2789	1.3310	1.8458
Logit	0.3730	0.2990	0.3730	0.2990	1.4129	1.7257
Probit	0.3730	0.2989	0.3730	0.2989	1.4123	1.7264

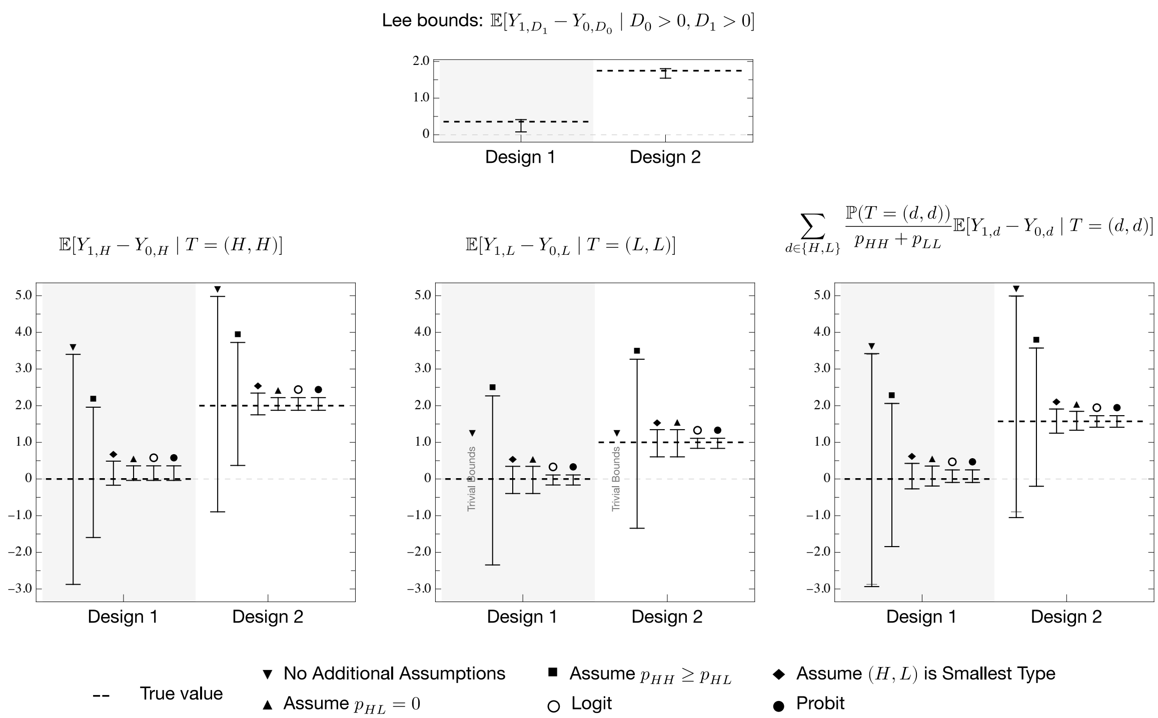
Notes: Outcome is simulated log earnings. p^*_t is the optimal value of p_t over the identified set for response-types under the given assumption; p^*_t may differ for lower and upper bounds and is therefore presented separately.

Our bounds on the within-firm effects $\mathbb{E}(Y_{1,H}-Y_{0,H}|T=(H,H))$ and $\mathbb{E}(Y_{1,L}-Y_{0,L}|T=(L,L))$ always include 0 across the different scenarios (which correspond to different assumptions on response types).

In design 2, Lee bounds also lie entirely within the positive quadrant, excluding 0. However, in this case, the DGP is consistent with the true within-firm effects being strictly positive. Interestingly, our bounds for $\mathbb{E}(Y_{1,H}-Y_{0,H}|T=(H,H))$ and $\mathbb{E}(Y_{1,L}-Y_{0,L}|T=(L,L))$ lie entirely within the positive quadrant, excluding 0, when restricting the response types, showing that they are informative enough to reveal the effect of job training on wages.

Appendix D. Job Corps Study: Additional Empirical Tables and Figures for Job Corps RCT

This section provides institutional background and summary statistics for the Job Corps program, alongside the empirical evidence of treatment-induced sorting into high-type (amenity-providing) firms referenced in the main text. Additionally, it details our replication of Lee's bounds and then presents the identified sets for response



Notes: Dashed black line indicates the true value of the parameter. In all panels, the solid black lines indicate sharp bounds; in the final panel, the gray lines indicate the (generally invalid) result of the "naïve approach" of taking the weighted average of firm-level bounds (see Remark 4).

FIGURE 3. Lee (2009) Bounds and Multilayered Bounds in Simulations.

types for weeks 135, 180 and 208 of the Job Corps study. Finally, it provides tabular counterparts to the bounds figures displayed in the main text.

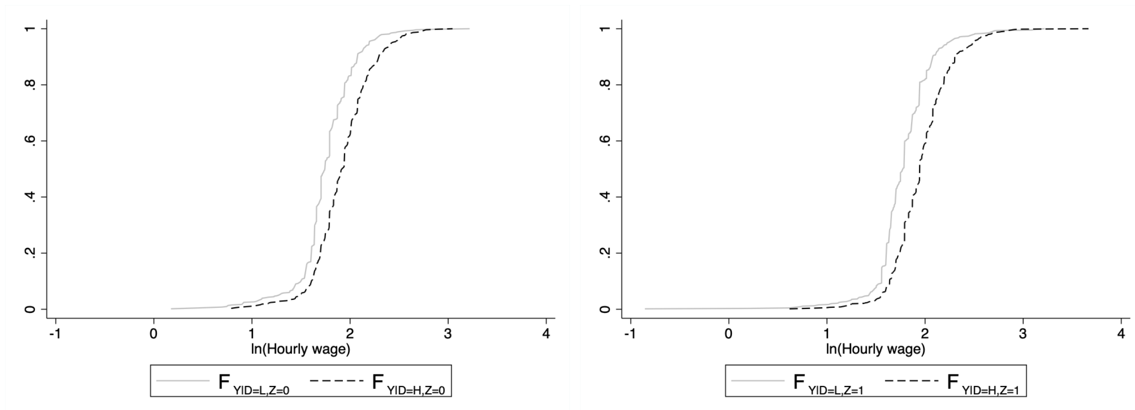
D.1. **Job Corps program.** Job Corps is the largest residential career training program in the U.S. It is free for participants and targets disadvantaged people aged 16 to 24 with the aim of helping these people become more responsible, employable, and productive citizens (Johnson et al. 1999). Most participants live at a local Job Corps center and complete 440 hours of academic instruction and 700 hours of vocational training. Job Corps also provides job search assistance upon participant completion of the program. The typical participant completes the program over a span of 30 weeks. Job Corps has trained more than two million individuals since its inception under the Economic Opportunity Act of 1964. The program trains over 60,000 enrollees per year, at roughly 130 Job Corps centers nationwide, with an estimated cost of 34,301 USD per enrollee and 57,312 USD per graduate (Liu et al. 2020).[3]

D.2. **Job Corps Study.** During the mid- to late 1990s, the U.S. Department of Labor funded a randomized evaluation of Job Corps, implemented by Mathematica Policy Research, Inc. Existing evaluations of the Job Corps Study include Schochet et al. (2001), Schochet et al. (2008), Lee (2009), and Blanco et al. (2013). The Job Corps Study randomized 80,883 eligible individuals who applied to Job Corps for the first time between November 1994 and December 1995 into two groups: (i) 5,977 individuals into the control group (embargoed from participating in Job Corps for three years) and (ii) 74,906 individuals into the treatment group. Of the 74,906 individuals assigned to treatment, 9,409 were randomly selected for data collection. All control individuals were selected for data collection. The final sample consists of 15,386 participants who were interviewed at the time of random assignment and then subsequently 12, 30, and 48 months after random assignment.

D.3. **Summary Statistics.** Table 5 presents summary statistics for the sample of individuals who have non-missing values for weekly earnings and hours for every week following random assignment. Means and standard deviations for a number of baseline and post-randomization variables are reported separately by treatment status. Consistent with successful randomization in the National Job Corps Study, the table shows that there are no statistical differences in the means of demographic, education, background and baseline employment/income variables across treatment and control groups. Table 5 also shows economically and statistically significant differences in employment and earnings outcomes by treatment status, post randomization. After 208 weeks, the hours and earnings of the treatment group are approximately 8% and 14% higher, respectively, than those of the control group.

[3]Included in these costs are direct transfers to Job Corps participants. The average participant in the 1995 Job Corps randomized evaluation (detailed below) received 2,361 USD in direct transfers, which consisted of 1,427 USD in pay, 704 USD of food and 230 USD of clothing; these transfers accounted for approximately 14% of the cost per participant at the time (McConnell and Glazerman 2001).

D.4. **Joint Distribution of Wages and Amenities.** Figure 4 presents the empirical cumulative distribution functions of log wage by firm type, classifying firms according to the provision of health insurance, for treatment and control groups at week 90.[4,5] For both treated and control units, the distribution of log wages for firms that provide health insurance stochastically dominates the distribution of log wages for firms that do not.[6] This suggests that firms providing amenities pay higher wages than firms that do not.[7]



(a) control units (b) treated units

Notes: Hourly wage calculated as weekly earnings divided by weekly hours for the employed.

FIGURE 4. Cumulative distribution function by firm type (provision of health insurance) at week 90. RCT = Job Corps Study.

---

[4]At the time of the National Job Corps study, there were no legal requirements for firms to provide health insurance and, conditional on firm provision, federal law generally prohibited discriminatory provision across workers (United States Equal Employment Opportunity Commission 2009). The relevant federal laws at the time of the National Job Corps Study included the following. Title VII of the Civil Rights Act of 1964; the Age Discrimination in Employment Act of 1967; Title I and Title V of the Americans with Disabilities Act of 1990 (United States Equal Employment Opportunity Commission 2009).

[5]We present results for week 90, following preferred specification of Lee (2009); results for alternative weeks are available upon request.

[6]At week 90, mean wages at firms that offered amenities were approximately 15% higher compared to firms that did not.

[7]Of course, it is possible that firms pay compensating differentials which causally reduce wages. The evidence presented here shows that the cross-sectional variation across firms dominates the variation within firms. This is consistent with evidence in Lamadon et al. (2022) who show that high-amenities firms are also more productive firms.

D.5. **Differential Sorting of Treatment and Control Workers.** Table 6 presents the probability of working in a firm (conditional on employment) that provides observable amenities, at week 90, according to the status of treatment. The evidence shows that treated individuals are more likely to work at firms with job amenities in all but one case (and we also find that this trend persists across all weeks). This is consistent with the evidence presented in Schochet et al. (2008).

D.6. **Replication of Lee (2009) Bounds.** Table 7 reports our replication of the bounds reported in Lee (2009) for weeks 90, 135, 180 and 208. We do not report bounds for week 45 since we discovered that the monotonicity assumption is violated. Table 7 reports Lee's bounds when treating ln(hourly wage) as a continuous variable (as we do throughout the paper). All quantities are very close to the estimates in Lee (2009), though a small difference arises in the bounds due to Lee's use of vingtiles of ln(hourly wage). Table 8 shows that when we use vingtiles of ln(hourly wage), the bounds are identical to the ones reported in Lee (2009).

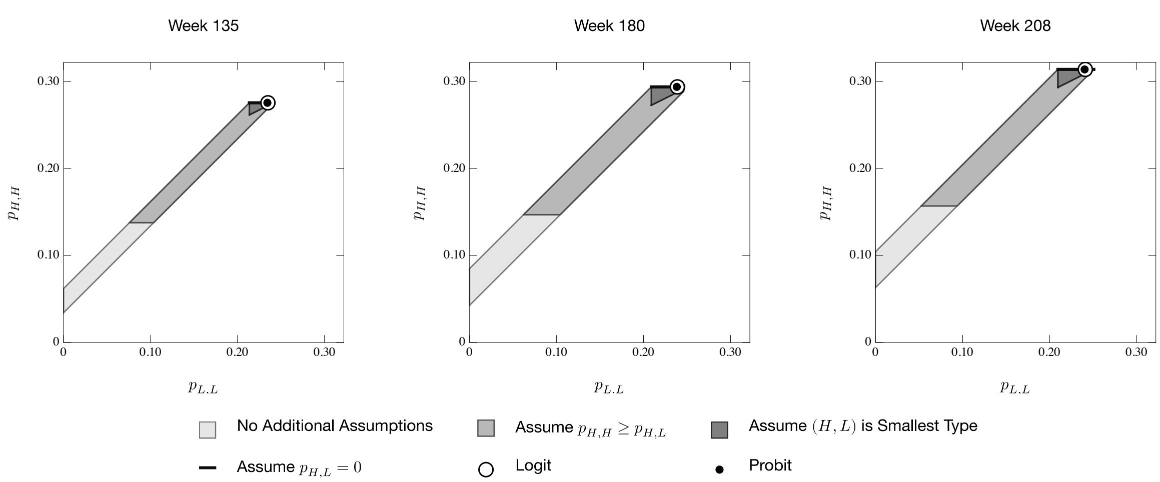


FIGURE 5. Identified set for $(p_{L,L}, p_{H,H})$ at weeks 135, 180 and 208. RCT = Job Corps Study.

Appendix E. Description of WorkAdvance RCTs and Additional Empirical Tables and Figures

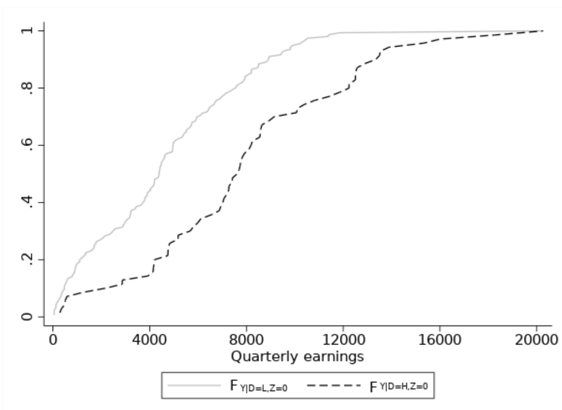
This section expands on the WorkAdvance RCTs by providing additional information on program curricula and summary statistics, alongside the empirical evidence of treatment-induced sorting into high-type (target-sector) firms referenced in the main text. Finally, it presents tabular counterparts to the bounds figures displayed in the main text.

E.1. **WorkAdvance Program and RCTs.** As detailed in the main text, our analysis focuses on the Madison Strategies and Towards Employment WorkAdvance RCTs. Although they targeted different in-demand sectors, the two programs shared a highly similar core structure. Both ranged from 4 to 32 weeks in duration, with curricula centered on pre-employment career readiness and sector-specific skills training. Each provider also integrated sector-specific job placement assistance with post-employment retention and advancement services. Upon completion of both programs, participants received the appropriate certification. While both evaluations targeted low-income adults, their screening requirements differed slightly: Madison Strategies required applicants to test at an eighth-grade level, pass a behavioral assessment, pass mechanical aptitude and manual dexterity exams, and hold a valid driver's license; by contrast, Towards Employment required testing at a sixth- to tenth-grade level alongside passing background and drug screenings.

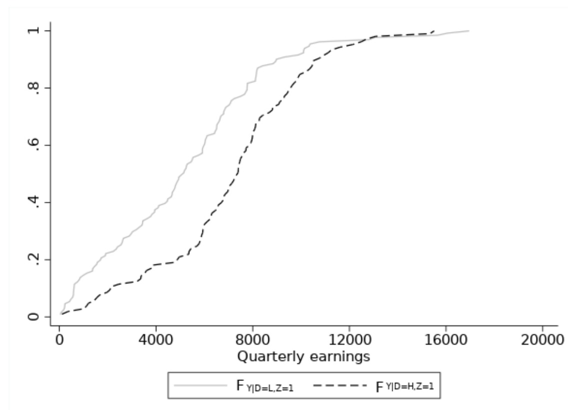
E.2. **Summary Statistics.** Tables 11 and 12 present summary statistics for the Madison Strategies and Towards Employment RCTs, respectively. Means and standard deviations for a number of baseline and post-randomization variables are reported separately by treatment status. Consistent with successful randomization, there are no statistical differences in the means of demographic, education, and baseline employment/income variables across treatment and control groups.

E.3. **Joint Distribution of Wages and Amenities.** Figure 6 presents the empirical cumulative distribution functions of quarterly earnings for the treatment and control groups, with firms classified by whether they operate within the target sector. Complementing our earlier evidence for amenity-providing firms in the Job Corps study, the log wage distribution for target-sector firms stochastically dominates that of non-target-sector firms across both treatment and control groups, in both WorkAdvance RCTs.

RCT = Madison Strategies

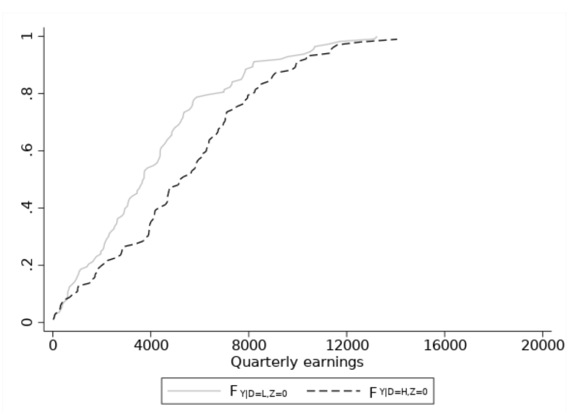


(a) control units

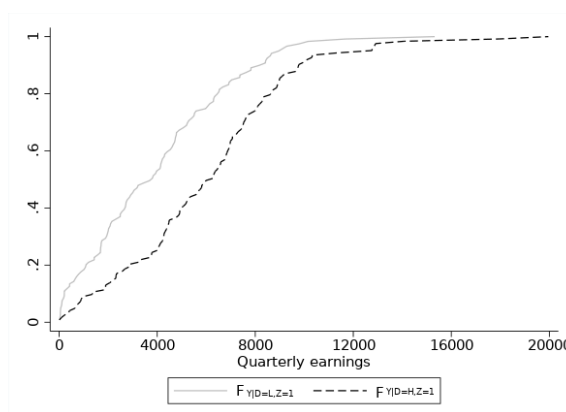


(b) treated units

RCT = Towards Employment



(c) control units



(d) treated units

Notes: Quarterly earnings are for the employed (i.e., does not include 0s for the unemployed).

Figure 6. Cumulative distribution function by firm sector. RCT = WorkAdvance.

	Control		Treated		Difference	
	Mean	S.D.	Mean	S.D.	Difference	S.E.
Female	0.46	0.50	0.45	0.50	-0.01	0.01
Age at baseline	18.35	2.10	18.44	2.16	0.09	0.05
White, non-Hispanic	0.26	0.44	0.27	0.44	0.00	0.01
Black, non-Hispanic	0.49	0.50	0.49	0.50	0.00	0.01
Hispanic	0.17	0.38	0.17	0.37	-0.00	0.01
Other race/ethnicity	0.07	0.26	0.07	0.26	-0.00	0.01
Never married	0.92	0.28	0.92	0.28	0.00	0.01
Married	0.02	0.15	0.02	0.14	-0.00	0.00
Living together	0.04	0.20	0.04	0.19	-0.00	0.00
Separated	0.02	0.14	0.02	0.15	0.00	0.00
Has child	0.19	0.39	0.19	0.39	-0.00	0.01
Number of children	0.27	0.64	0.27	0.65	0.00	0.01
Education	10.11	1.54	10.11	1.56	0.01	0.03
Mother's education	11.46	2.59	11.48	2.56	0.02	0.06
Father's education	11.54	2.79	11.39	2.85	-0.15	0.08
Ever arrested	0.25	0.43	0.25	0.43	-0.00	0.01
Household income						
<3,000	0.25	0.43	0.25	0.44	0.00	0.01
3,000-6,000	0.21	0.41	0.21	0.40	-0.00	0.01
6,000-9,000	0.11	0.32	0.12	0.32	0.00	0.01
9,000-18,000	0.24	0.43	0.24	0.43	-0.00	0.01
>18,000	0.18	0.39	0.18	0.38	-0.00	0.01
Personal income						
<3,000	0.79	0.41	0.79	0.41	-0.00	0.01
3,000-6,000	0.13	0.34	0.13	0.33	-0.00	0.01
6,000-9,000	0.05	0.21	0.05	0.22	0.01	0.00
>9,000	0.03	0.18	0.03	0.17	-0.00	0.00
At baseline						
Have job	0.19	0.39	0.20	0.40	0.01	0.01
Mths. empl. prev. yr.	3.53	4.24	3.60	4.25	0.07	0.09
Had job, prev. yr.	0.63	0.48	0.63	0.48	0.01	0.01
Earnings, prev. yr.	2810.48	4435.62	2906.45	6401.33	95.97	117.10
Usual hours/week	20.91	20.70	21.82	21.05	0.91	0.45
Usual weekly earn.	102.89	116.46	110.99	350.61	8.10	5.09
Post randomization						
Week 52 hours	17.78	23.39	15.30	22.68	-2.49	0.49
Week 104 hours	21.98	26.08	22.64	26.25	0.67	0.56
Week 156 hours	23.88	26.15	25.88	26.57	2.00	0.56
Week 208 hours	25.83	26.25	27.79	25.74	1.95	0.56
Week 52 earn.	103.80	159.89	91.55	149.28	-12.25	3.33
Week 104 earn.	150.41	210.24	157.42	200.27	7.02	4.42
Week 156 earn.	180.88	224.43	203.71	239.80	22.84	4.94
Week 208 earn.	200.50	230.66	227.91	250.22	27.41	5.11
Total 4 yr. earn.	30006.69	26893.60	30800.41	26437.39	793.72	571.83
Sample size	3599		5546		9145	

Notes: Weekly earnings calculated as the sum of total earnings in a given week and are not conditional on employment (i.e., includes 0s for the unemployed).

TABLE 5. Summary statistics by treatment status. RCT = Job Corps Study.

	Control		Treated		Difference	
	Mean	S.D.	Mean	S.D.	Difference	S.E.
Health insurance	0.4860	0.5000	0.5072	0.5001	0.0211	0.0186
Paid sick leave	0.3922	0.4884	0.4251	0.4945	0.0329	0.0183
Paid vacation	0.5407	0.4985	0.5800	0.4937	0.0393	0.0180
Childcare assistance	0.1311	0.3376	0.1422	0.3494	0.0111	0.0127
Flexible hours	0.5330	0.4991	0.5568	0.4969	0.0237	0.0183
Employer-provided transportation	0.1973	0.3981	0.1876	0.3905	-0.0097	0.0147
Pension or retirement benefits	0.3670	0.4822	0.3938	0.4887	0.0268	0.0180
Dental plan	0.3863	0.4871	0.4288	0.4950	0.0425	0.0182
Tuition reimbursement	0.2212	0.4152	0.2602	0.4389	0.0390	0.0158
Employment	0.4368	0.4961	0.4391	0.4963	0.0022	0.0111
Sample size	3288		5091		8379	

Notes: Control and treatment probabilities have interpretation as $P[D = H|D > 0, Z = z]$ for $z \in \{0, 1\}$, respectively, when classifying firms as type H if they provide a given amenity.

TABLE 6. Probability of working at amenity-providing firm at week 90 (conditional on employment). RCT = Job Corps Study.

	$\mathbb{P}[D > 0\|Z = 0]$	$\mathbb{P}[D > 0\|Z = 1]$	Trimming Proportion	$\mathbb{E}[Y_{1,D_1} - Y_{0,D_0}\|D_0 > 0, D_1 > 0]$ lower	upper
Week 90	0.4600	0.4601	0.0003	0.0468	0.0484
Week 135	0.5173	0.5451	0.0509	-0.0072	0.0842
Week 180	0.5403	0.5825	0.0724	-0.0325	0.0901
Week 208	0.5655	0.6068	0.0680	-0.0217	0.0989

Notes: Treatment bounds are for ln(hourly wage), where hourly wage equals weekly earnings divided by weekly hours for the employed. Propensity scores and trimming proportions are numerically equivalent to Lee (2009). The trimming proportion, which Lee reports, is equal to $1 - p$ where $p \equiv \mathbb{P}(AE)$, e.g., the share of always-employed among employed individuals receiving job training. The slight numerical difference in bounds arises as Lee uses vingtiles of ln(hourly wage); these are presented in Table 8.

TABLE 7. Lee's bounds: continuous ln(hourly wage). RCT = Job Corps Study.

	$\mathbb{P}[D > 0\|Z = 0]$	$\mathbb{P}[D > 0\|Z = 1]$	Trimming Proportion	$\mathbb{E}[Y_{1,D_1} - Y_{0,D_0}\|D_0 > 0, D_1 > 0]$ lower	upper
Week 90	0.4600	0.4601	0.0003	0.0423	0.0428
Week 135	0.5173	0.5451	0.0509	-0.0159	0.0757
Week 180	0.5403	0.5825	0.0724	-0.0325	0.0868
Week 208	0.5655	0.6068	0.0680	-0.0194	0.0933

Notes: Treatment bounds are for vingtiles of ln(hourly wage), where hourly wage equals weekly earnings divided by weekly hours for the employed. Propensity scores and trimming proportions are numerically equivalent to Lee (2009). The trimming proportion, which Lee reports, is equal to $1 - p$ where $p \equiv \mathbb{P}(AE)$, e.g., the share of always-employed among employed individuals receiving job training.

TABLE 8. Lee's bounds: vingtiles of ln(hourly wage). RCT = Job Corps Study.

			$\mathbb{E}(Y_{1,H} - Y_{0,H} \mid T = (H,H))$		$\mathbb{E}(Y_{1,L} - Y_{0,L} \mid T = (L,L))$	
Week 90	$p^*_{H,H}$	$p^*_{L,L}$	lower	upper	lower	upper
Baseline	0.0010	0.0000	-2.1415	2.3907		
$p_{H,H} \geq p_{H,L}$	0.1120	0.1108	-0.4214	0.5020	-0.4002	0.4542
(H,L) is smallest type	0.2239	0.2228	-0.0023	0.0754	-0.0191	0.0673
$p_{H,L} = 0$	0.2239	0.2228	-0.0018	0.0750	-0.0191	0.0673
Logit	0.2239	0.2229	-0.0018	0.0750	-0.0180	0.0656
Probit	0.2239	0.2229	-0.0018	0.0750	-0.0180	0.0656
Week 135						
Baseline	0.0344	0.0000	-1.1228	1.1454		
$p_{H,H} \geq p_{H,L}$	0.1379	0.0758	-0.5075	0.5433	-0.6064	0.7090
(H,L) is smallest type	0.2619	0.2137	-0.1135	0.1369	-0.1180	0.1672
$p_{H,L} = 0$	0.2758	0.2137	-0.0529	0.0732	-0.1180	0.1672
Logit	0.2758	0.2349	-0.0529	0.0732	-0.0326	0.0847
Probit	0.2758	0.2343	-0.0529	0.0732	-0.0367	0.0881
Week 180						
Baseline	0.0429	0.0000	-1.0454	1.1410		
$p_{H,H} \geq p_{H,L}$	0.1471	0.0620	-0.5063	0.5525	-0.7710	0.8503
(H,L) is smallest type	0.2730	0.2090	-0.1421	0.1704	-0.1806	0.2110
$p_{H,L} = 0$	0.2941	0.2090	-0.0552	0.0854	-0.1806	0.2110
Logit	0.2941	0.2389	-0.0552	0.0854	-0.0675	0.1020
Probit	0.2941	0.2382	-0.0552	0.0854	-0.0712	0.1056
Week 208						
Baseline	0.0633	0.0000	-0.8821	0.9895		
$p_{H,H} \geq p_{H,L}$	0.1571	0.0525	-0.4888	0.5670	-0.9059	0.8730
(H,L) is smallest type	0.2935	0.2096	-0.1217	0.1797	-0.1914	0.2082
$p_{H,L} = 0$	0.3142	0.2096	-0.0430	0.1016	-0.1914	0.2082
Logit	0.3142	0.2411	-0.0430	0.1016	-0.0771	0.0952
Probit	0.3142	0.2403	-0.0430	0.1016	-0.0809	0.0995

Notes: Treatment bounds are for ln(hourly wage); hourly wage calculated as weekly earnings divided by weekly hours for the employed. p^*_t is the minimum value of p_t over $\Theta_I(\mathcal{R}_T)$, for the corresponding $\mathcal{R}_T$.

TABLE 9. Multilayered bounds based on health insurance amenity. RCT = Job Corps Study.

	p^*_{HH}	p^*_{LL}	p^*_{HH}	p^*_{LL}	$\sum_{d\in\{L,H\}} \frac{p_{d,d}}{p_{HH}+p_{LL}} \mathbb{E}[Y_{1,d} - Y_{0,d} \mid T=(d,d)]$	
Week 90	(for lower bound)		(for upper bound)		lower	upper
Baseline	0.0014	0.0004	0.0014	0.0004	-2.4634	2.4588
$p_{H,H} \geq p_{H,L}$	0.1120	0.1109	0.1120	0.1109	-0.4106	0.4780
(H,L) is smallest type	0.2235	0.2224	0.2235	0.2224	-0.0136	0.0749
$p_{H,L} = 0$	0.2235	0.2224	0.2235	0.2224	-0.0136	0.0749
Logit	0.2239	0.2229	0.2239	0.2229	-0.0099	0.0703
Probit	0.2239	0.2229	0.2239	0.2229	-0.0099	0.0703
Week 135						
Baseline	0.0366	0.0022	0.0392	0.0048	-1.1602	1.2762
$p_{H,H} \geq p_{H,L}$	0.1380	0.0758	0.1380	0.0758	-0.5423	0.6017
(H,L) is smallest type	0.2614	0.2127	0.2614	0.2127	-0.1179	0.1532
$p_{H,L} = 0$	0.2754	0.2132	0.2754	0.2132	-0.0833	0.1165
Logit	0.2758	0.2349	0.2758	0.2349	-0.0436	0.0785
Probit	0.2758	0.2343	0.2758	0.2343	-0.0454	0.0800
Week 180						
Baseline	0.0475	0.0046	0.0485	0.0056	-1.1369	1.2631
$p_{H,H} \geq p_{H,L}$	0.1471	0.0620	0.1471	0.0620	-0.5846	0.6407
(H,L) is smallest type	0.2727	0.2081	0.2727	0.2081	-0.1607	0.1900
$p_{H,L} = 0$	0.2937	0.2086	0.2937	0.2086	-0.1101	0.1397
Logit	0.2941	0.2389	0.2941	0.2389	-0.0607	0.0929
Probit	0.2941	0.2382	0.2941	0.2382	-0.0624	0.0945
Week 208						
Baseline	0.0705	0.0072	0.0671	0.0038	-0.9987	1.0704
$p_{H,H} \geq p_{H,L}$	0.1571	0.0525	0.1571	0.0525	-0.5932	0.6436
(H,L) is smallest type	0.2928	0.2087	0.2928	0.2087	-0.1532	0.1941
$p_{H,L} = 0$	0.3138	0.2092	0.3138	0.2092	-0.1049	0.1461
Logit	0.3142	0.2411	0.3142	0.2411	-0.0578	0.0988
Probit	0.3142	0.2403	0.3142	0.2403	-0.0594	0.1007

Notes: Treatment bounds are for ln(hourly wage); hourly wage calculated as weekly earnings divided by weekly hours for the employed. p^*_t is the optimal value of p_t over the joint identified set for response-types under the given assumption; p^*_t may differ for lower and upper bounds and is therefore presented separately. Because p^*_t values are determined by a grid search, the values in columns 2-4 should be read only as the grid point used in the numerical approximation, not as an exact feasible response-type probability.

Table 10. Aggregate multilayered bounds based on health insurance amenity. RCT = Job Corps Study.

	Control		Treated		Difference	
	Mean	S.D.	Mean	S.D.	Difference	S.E.
Female	0.14	0.34	0.18	0.39	0.05	0.03
Adults $\leq$ 24	0.22	0.42	0.23	0.42	0.01	0.03
Black	0.35	0.48	0.39	0.49	0.03	0.04
HS/GED or less	0.40	0.49	0.43	0.50	0.03	0.04
At baseline:						
Have job	0.28	0.45	0.26	0.44	-0.02	0.03
Quarterly earnings	2257.07	2820.54	1944.51	2495.57	-312.56	201.91
8 quarters post randomization:						
Employment	0.66	0.48	0.67	0.47	0.01	0.04
Share employment in target sector	0.31	0.46	0.44	0.50	0.14	0.04
Quarterly earnings	3703.87	4169.10	4049.91	4003.27	346.04	309.72
Sample size	344		353		697	

Notes: Quarterly earnings are not conditional on employment (i.e., includes 0s for the unemployed).

TABLE 11. Summ. stats. by treatment status, RCT: Madison Strategies

	Control		Treated		Difference	
	Mean	S.D.	Mean	S.D.	Difference	S.E.
Female	0.59	0.49	0.58	0.49	-0.01	0.04
Adults $\leq$ 24	0.23	0.42	0.22	0.42	-0.01	0.03
Black	0.73	0.45	0.77	0.42	0.05	0.03
HS/GED or less	0.42	0.49	0.44	0.50	0.03	0.04
At baseline:						
Have job	0.27	0.45	0.26	0.44	-0.01	0.03
Quarterly earnings	1504.59	2201.72	1723.72	2452.60	219.12	176.42
8 quarters post randomization:						
Employment	0.62	0.49	0.69	0.46	0.08	0.04
Share employment in target sector	0.47	0.50	0.51	0.50	0.03	0.05
Quarterly earnings	3040.57	4024.21	3497.54	3684.01	456.97	292.04
Sample size	349		349		698	

Notes: Quarterly earnings are not conditional on employment (i.e., includes 0s for the unemployed).

TABLE 12. Summ. stats. by treatment status, RCT: Towards Employment

				$\mathbb{E}[Y_{1,D_1} - Y_{0,D_0} \| D_0 > 0, D_1 > 0]$	
	$\mathbb{P}[D > 0 \| Z = 0]$	$\mathbb{P}[D > 0 \| Z = 1]$	Trimming Proportion	lower	upper
Madison Strategies	0.6570	0.6686	0.0173	244.1928	524.9793
Towards Employment	0.6160	0.6934	0.1116	-676.8651	705.6605

Notes: Treatment bounds are for quarterly wages. The trimming proportion, which Lee (2009) reports, is equal to $1 - p$ where $p \equiv \mathbb{P}(AE)$, e.g., the share of always-employed among employed individuals receiving job training.

TABLE 13. Lee's bounds. RCTs = WorkAdvance.

			$\mathbb{E}(Y_{1,H} - Y_{0,H} \mid T = (H,H))$		$\mathbb{E}(Y_{1,L} - Y_{0,L} \mid T = (L,L))$	
Madison Strategies	$p^*_{H,H}$	$p^*_{L,L}$	lower	upper	lower	upper
Baseline	0.0000	0.1560			-6172.4634	7139.9806
$p_{H,H} \geq p_{H,L}$	0.1017	0.2578	-7380.1460	5628.8654	-3353.0004	4562.1007
(H,L) is smallest type	0.1977	0.3595	-2652.6847	1216.5151	-798.7005	2016.2079
$p_{H,L} = 0$	0.2035	0.3595	-2353.4366	806.8709	-798.7005	2016.2079
Logit	0.2035	0.3711	-2353.4366	806.8709	-328.9279	1718.8632
Probit	0.2035	0.3711	-2353.4366	806.8709	-328.9279	1718.8632
Towards Employment						
Baseline	0.0000	0.0000				
$p_{H,H} \geq p_{H,L}$	0.1461	0.1175	-5784.7451	6488.1176	-6510.9268	5975.5610
(H,L) is smallest type	0.2536	0.2636	-2044.9605	3064.8644	-2335.5870	1932.3913
$p_{H,L} = 0$	0.2923	0.2636	-740.4216	1356.9608	-2335.5870	1932.3913
Logit	0.2923	0.3106	-740.4216	1356.9608	-1006.6086	464.6827
Probit	0.2923	0.3103	-740.4216	1356.9608	-1014.3705	474.0062

Notes: Outcome is quarterly wages. p^*_t is the optimal value of p_t over the identified set for response-types under the given assumption.

TABLE 14. Multilayered bounds. RCTs = WorkAdvance.

	p^*_{HH}	p^*_{LL}	p^*_{HH}	p^*_{LL}	$\sum_{d \in \{L,H\}} \frac{p_{d,d}}{p_{HH} + p_{LL}} \mathbb{E}[Y_{1,d} - Y_{0,d} \mid T = (d,d)]$	
Madison Strategies	(for lower bound)		(for upper bound)		lower	upper
Baseline	0.0145	0.1710	0.0090	0.1655	-6498.0976	7255.9812
$p_{H,H} \geq p_{H,L}$	0.1010	0.2570	0.1010	0.2570	-4514.3517	4886.1844
(H,L) is smallest type	0.1960	0.3545	0.1960	0.3545	-1593.0171	1848.4113
$p_{H,L} = 0$	0.2010	0.3570	0.2010	0.3570	-1459.5360	1688.9594
Logit	0.2035	0.3711	0.2035	0.3711	-1045.8942	1395.8872
Probit	0.2035	0.3711	0.2035	0.3711	-1045.8942	1395.8872
Towards Employment						
Baseline	0.0000	0.0000	0.0000	0.0000		
$p_{H,H} \geq p_{H,L}$	0.1450	0.1170	0.1450	0.1170	-6142.9166	6288.9131
(H,L) is smallest type	0.2500	0.2590	0.2500	0.2590	-2318.5475	2613.7796
$p_{H,L} = 0$	0.2900	0.2620	0.2900	0.2620	-1562.2141	1840.9057
Logit	0.2923	0.3106	0.2923	0.3106	-877.5657	897.2439
Probit	0.2923	0.3103	0.2923	0.3103	-881.5065	902.2352

Notes: Outcome is quarterly wages. p^*_t is the optimal value of p_t over the joint identified set for response-types p^*_t under the given assumption; p^*_t may differ for lower and upper bounds and is therefore presented separately. Because p^*_t values are determined by a grid search, the values in columns 2-4 should be read only as the grid point used in the numerical approximation, not as an exact feasible response-type probability.

TABLE 15. Aggregate multilayered bounds. RCTs = WorkAdvance

Appendix F. 'Top 5' papers with multilayered sample selection

In the literature survey we conducted, referenced in the introduction, we counted 56 papers published in 'top 5' general interest economic journals that cited Lee (2009) and 42 that empirically implemented Lee bounds.[8] This section details 6 of these papers that feature multilayered selection, where researchers simplified the sample selection problem by collapsing it to a single dimension.

Daruich et al. (2023) Studies a 2001 Italian reform which lifted constraints on the employment of temporary contract workers but maintained employment protection laws for permanent contract employees. The outcome is individual earnings. Lee bounds are employed to address concern that the reform affected employment. Conditional on labor market entry, the reform can affect worker sorting across industries and/or firms; indeed the paper finds meaningful changes in the shares of temporary contract workers in certain industries.

Cullen and Pakzad-Hurson (2023) Studies state-level laws in the U.S. protecting the right of private sector workers to discuss salary information with co-workers. The main outcome is worker wages. They estimate Lee bounds to address sample selection into employment. The treatment may also affect worker sorting across firms; for example, knowledge of co-workers' salaries could cause workers to sort to firms with flatter pay hierarchies.

Bianchi and Giorcelli (2022) Studies the impact of the Training Within Industry program, a U.S. government training program intended to be provided to all firms involved in war production between 1940 and 1945. The main outcome is firm total factor productivity. They estimate Lee bounds to address the higher attrition rate of untrained firms: treated firms had 90% survival rate at least 10 years following treatment whereas control firms only had 64% survival rate. Conditional on firm survival, training may also affect the sorting of firms across industries or other important dimensions. This is particularly plausible in this paper's setting where: (i) trained firms undertook structural changes transforming them into larger and more complex

[8]The 'top 5' refer to the American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics and the Review of Economic Studies.

organizations; (ii) this paper estimates treatment bounds for outcomes observed post-Second World War, when many firms would have plausibly switched industries (i.e., left war production).

Fink et al. (2020) Evaluate an experiment offering subsidized loans to randomly selected villages in rural Zambia where farmers suffer from liquidity constraints in the months prior to harvest (i.e., the lean season). The outcomes are individual- and village-level earnings. They estimate Lee bounds to address that the likelihood of entering the labor market decreases with the loan treatment. Conditional on entry to the labor market, treatment has potential to affect the types of jobs individuals accept; this is particularly plausible in this paper's setting as labor sales occur within villages between better- and worse-off farmers at individually negotiated rates.

Giorcelli (2019) Studies long-run effects of U.S. Technical Assistance and Productivity Program (USTAPP) which provided management training and technologically advanced machines to Italian firms from 1952 to 1958. Outcomes include firm-level sales, number of employees and total factor productivity revenue. They estimate Lee bounds to address the treatment-control difference in firm survival probability. As in Bianchi and Giorcelli (2022), conditional on firm survival, management training likely affects the sorting of firms across industries.

Fisman et al. (2017) Estimate effect of cultural proximity on loan outcomes for lenders and borrowers using dyadic data on religion and caste for lending officers and borrowers from a state-owned Indian bank. Outcomes include amount of debt received, total number of borrowers and average loan size. They estimate Lee bounds as outcomes are only observed conditional on a group receiving credit. Conditional on a group receiving credit, "same group matches" also plausibly affect the type of loans a group receives. For example, "same group match" borrowers may receive favorable loan terms; this paper indeed notes the potential for these effects but is constrained by data limitations.

References

Bianchi, Nicola and Michela Giorcelli (2022): "The Dynamics and Spillovers of Management Interventions: Evidence from the Training within Industry Program," *Journal of Political Economy*, 130, 1630–1675.

Blanco, German, Carlos A Flores, and Alfonso Flores-Lagunes (2013): "The Effects of Job Corps Training on Wages of Adolescents and Young Adults," *American Economic Review*, 103, 418–422.

Cullen, Zoë B. and Bobak Pakzad-Hurson (2023): "Equilibrium Effects of Pay Transparency," *Econometrica : journal of the Econometric Society*, 91, 765–802.

Daruich, Diego, Sabrina Di Addario, and Raffaele Saggio (2023): "The Effects of Partial Employment Protection Reforms: Evidence from Italy," *Review of Economic Studies*, 90, 2880–2942.

Fink, Günther, B. Kelsey Jack, and Felix Masiye (2020): "Seasonal Liquidity, Rural Labor Markets, and Agricultural Production," *American Economic Review*, 110, 3351–3392.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig (2017): "Cultural Proximity and Loan Outcomes," *American Economic Review*, 107, 457–492.

Giorcelli, Michela (2019): "The Long-Term Effects of Management and Technology Transfers," *American Economic Review*, 109, 121–152.

Johnson, Terry, Mark Gritz, Russell Jackson, John Burghardt, Carol Boussy, Jan Leonard, and Carlyn Orians (1999): "National Job Corps Study: Report on the Process Analysis," Tech. rep., Mathematica Policy Research, Princeton, NJ.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2022): "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 112, 169–212.

Lee, David S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102.

Liu, Albert, Clayton Perry, Robbie Skinner, and Scott Cody (2020): "Estimating Job Corps Cost per Enrollee and Cost per Graduate," Tech. rep., Insight Policy Research, Arlington, VA.

McConnell, Sheena and Steven Glazerman (2001): "National Job Corps Study: The Benefits and Costs of Job Corps." Tech. Rep. ED457357, Mathematica Policy Research, Washington, DC.; Battelle Memorial Inst., Seattle, WA.; Decision Information Resources, Inc., Houston, TX.

Pearl, Judea (2001): "Direct and Indirect Effects," Tech. rep.

——— (2009): "Causal Inference in Statistics: An Overview," *Statistics Surveys*, 3.

Schochet, Peter, John Burghardt, and Steven Glazerman (2001): "National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes," Tech. rep., Mathematica Policy Research, Princeton, NJ.

Schochet, Peter Z, John Burghardt, and Sheena McConnell (2008): "Does Job Corps Work? Impact Findings from the National Job Corps Study," *American Economic Review*, 98, 1864–1886.

United States Equal Employment Opportunity Commission (2009): "Federal Laws Prohibiting Job Discrimination Questions and Answers," Tech. rep., United States Equal Employment Opportunity Commission, Washington, DC.