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

September, 2014

Abstract
Randomized and natural experiments are commonly used in economics and other social science fields to estimate the effect of programs and interventions. Even when employing experimental data, assessing the impact of a treatment is often complicated by the presence of sample selection (outcomes are only observed for a selected group) and noncompliance (some treatment group individuals do not receive the treatment while some control individuals do). We address both of these identification problems simultaneously and derive nonparametric bounds for average treatment effects within a principal stratification framework. We employ these bounds to empirically assess the wage effects of Job Corps (JC), the most comprehensive and largest federally-funded job training program for disadvantaged youth in the United States. Our results strongly suggest positive average effects of JC on wages for individuals who comply with their treatment assignment and would be employed whether or not they enrolled in JC (the “always-employed compliers”). Under relatively weak monotonicity and mean dominance assumptions, we find that this average effect is between 5.7 and 13.9 percent four years after randomization, and between 7.7 and 17.5 percent for Non-Hispanics. Our results are consistent with larger effects of JC on wages than those found without adjusting for noncompliance.

Key words and phrases: Instrumental variables; Nonparametric partial identification; Principal stratification; Program evaluation; Training programs
JEL classification: C13, C21, J30
1 Introduction

Randomized and natural experiments are commonly used in economics and other fields to estimate the effect of programs and interventions. Even when employing experimental data, assessing the impact of a treatment is often complicated by two critical identification problems: sample selection and noncompliance with the assigned treatment. Our leading example in the paper is the evaluation of the effect of the Job Corps (JC) training program on wages using data on individuals who were randomly assigned to participate or not in the program. In this case, the sample selection issue arises because wages are only observed for those who are employed, with the employment decision itself being potentially affected by the program. The noncompliance problem appears because some treatment group individuals did not enroll in the program, while some of the control individuals did enroll. In this paper, we derive nonparametric bounds for average treatment effects in settings where both identification problems are present, and employ these bounds to empirically assess the effect of JC on its participants’ wages.

JC is the most comprehensive and largest federally-funded job training program in the United States for disadvantaged youth between the ages of 16 and 24. It provides academic, vocational, and social skills training, among many other services (e.g., health care and job search assistance), at over 120 centers nationwide. A typical JC student lives at a local JC center for eight months and receives about 1,100 hours of academic and vocational instruction, which is equivalent to approximately one year of high school (Schochet et al., 2001). Assessing the effect of this and similar programs on wages is of great importance to policy makers. Most of the econometric evaluations of training programs, however, focus on their impact on total earnings, which are the product of the hourly wage and the hours worked. As discussed by Lee (2009), focusing only on total earnings fails to answer the relevant question of whether the programs lead to an increase in participants’ wages (e.g., through human capital accumulation), or to an increase in the probability of being employed (e.g., through job search assistance services) without any increase in wages.

Standard approaches for point identifying (i.e., theoretically learning the true parameter values in infinite samples) treatment effects in the presence of sample selection require strong
parametric assumptions or the availability of a valid instrument (e.g., Heckman, 1979). In settings where an instrument is unavailable, a recent literature has focused on bounding or partially identifying these effects under relatively mild assumptions. Part of this literature uses principal stratification (Frangakis and Rubin, 2002), which provides a framework for studying causal effects when controlling for a variable that has been affected by the treatment (in our setting, the employment decision). Principal stratification compares individuals within subpopulations (called principal strata) whose individuals share the same potential values of the employment variable under both treatment arms. In this spirit, Zhang, Rubin, and Mealli (2008) (hereafter ZRM) and Lee (2009) derive bounds for the average effect of a training program on wages for a particular stratum: the “always-employed” (those individuals who would be employed whether or not they were assigned to enroll in the program). They focus on this stratum because it is the only one for which the individuals’ wages are observed under both treatment arms. Following Zhang and Rubin (2003), ZRM consider two assumptions. The first, also used in Lee (2009), is a monotonicity assumption on the effect of the treatment (training program) on the selection (employment), and the second is a stochastic dominance assumption comparing the potential outcomes of the always-employed to those of other strata. Lee (2009) uses his bounds to evaluate the wage effects of JC, and Blanco et al. (2013) employ the bounds derived by ZRM and Lee (2009), and their extension to quantile treatment effects by Imai (2008) to study the wage effects of JC for different demographic groups.

Huber and Mellace (2013) and Lechner and Melly (2010) derive bounds for subpopulations other than the always-employed. Huber and Mellace (2013) construct bounds on the effects for two other strata (those who would be employed only if assigned to the treatment group, and those who would be employed only if assigned to the control group), as well as for the “selected” subpopulation (those whose wages are observed). While their assumptions are similar to those in ZRM and Lee (2009), additional assumptions are required (e.g., bounded support of the outcome) because bounds are constructed for strata and subpopulations for which the outcome is never observed under one of the treatment states. Lechner and Melly (2010) derive bounds for mean and quantile treatment effects for the “treated and selected” subpopulation (employed individuals who received training). Contrary to the previously described literature, they do not
use a principal stratification approach. Their bounds, however, are also based on monotonicity and stochastic dominance assumptions. Similar to Huber and Mellace (2013), they require an outcome with bounded support to partially identify average effects.

The previously discussed literature, with the exception of Lechner and Melly (2010), focuses on the intention-to-treat (ITT) effect of being offered to participate in the training program, as it compares potential outcomes according to the assigned treatment and ignores noncompliance. The noncompliance issue is important in practice. For example, in our data set, which is based on that used by Lee (2009) and Blanco et al. (2013), only 73.8 percent of the individuals assigned to the treatment group enrolled in JC, while 4.4 percent of the individuals assigned to the control group enrolled in JC in the four years after randomization.

In this paper, we extend the partial identification results in ZRM and Lee (2009) to account for noncompliance. Thus, we bound the effect of actual enrollment in the program, rather than the effect of being allowed to enroll in the program. Our approach to account for noncompliance is based on the works by Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996) (hereafter, AIR), which are a special case of principal stratification. They use an instrumental variable (IV) approach to address noncompliance in the absence of sample selection and show that, under some assumptions, IV estimators point identify the average treatment effect for those who comply with their treatment assignment (the compliers). We use principal stratification to address sample selection and noncompliance simultaneously, and derive bounds for the average effect of participating in a training program on wages for the stratum of always-employed compliers. This stratum consists of those who comply with their treatment assignment and would be employed whether or not they enrolled in the program. Analogous to the cases analyzed in Imbens and Angrist (1994), AIR, ZRM, and Lee (2009), this is the only stratum for which wages are observed for individuals who enrolled and did not enroll in the program. In our application, this is the largest stratum (about 40 percent of the population).

Principal stratification has often been used to address a single post-treatment complication. While it is our understanding that this is one of the first papers deriving bounds for treatment effects within this framework accounting for more than one identification problem, there are a few papers that employ this framework to point identify treatment effects in the presence of
multiple complications. For example, Mattei and Mealli (2007) address noncompliance, sample selection, and missing outcomes (dropout) in randomized experiments using a parametric Bayesian approach, and employ their methods to evaluate the effects of a new teaching program on breast self-examination. A related paper that is particularly relevant in our setting is the one by Frumento et al. (2012), who perform a likelihood-based analysis of the effects of JC on employment and wages simultaneously addressing noncompliance, sample selection, and missing outcomes due to non-response. They stratify the population based on the potential values of the compliance behavior and employment status to address noncompliance and sample selection, and to address the missing-outcome problem they assume that the probability that the outcome is missing for a given individual is random conditional on a set of observable characteristics. Under some parametric assumptions, they point identify the effect of JC on wages for the always-employed compliers. Our paper complements the work by Frumento et al. (2012) by constructing nonparametric bounds for the effect of JC on wages based on an alternative set of assumptions. In the empirical section of the paper, we also present results that account for missing values due to non-response by using weights constructed based on non-public use data that account for sample design and non-response.

The contribution of this paper is twofold. First, we add to the partial identification literature by deriving nonparametric bounds for average treatment effects in the presence of sample selection and noncompliance. More generally, our bounds can be used in settings where two identification problems are present (e.g., sample selection and an endogenous treatment) and there is a valid IV to address one of the problems. For example, when assessing the effect of military service on future health using the Vietnam-era draft lottery as an IV to address the endogeneity of the decision to serve in the military (e.g., Angrist et al., 2010), our results could be used to bound the average effect for those who enrolled in the military because of the draft lottery (compliers) and would live at the time the outcome is measured regardless of their veteran status. Second, we contribute to the literature on the evaluation of JC (Schochet et al., 2001; Schochet et al., 2008; Lee, 2009; Flores-Lagunes et al., 2010; Flores et al., 2012; Frumento et al., 2012; Blanco et al., 2013) by evaluating the effect of JC on its participants’ wages. Our results suggest greater positive average effects of JC on wages than those found without
adjusting for noncompliance in Lee (2009) and Blanco et al. (2013). Under our monotonicity and mean dominance assumptions, we find that the average effect of JC on wages for the always-employed compliers is between 5.7 and 13.9 percent four years after randomization, and between 7.7 and 17.5 percent for Non-Hispanics.

The paper is organized as follows. Section 2 presents the econometric framework and the partial identification results. Section 3 empirically analyzes the wage effects of JC. Section 4 presents a simulation study based on our data which analyzes the performance of our bounds, especially in cases when our assumptions are violated. Section 5 provides our conclusions.

2 Econometric Framework

2.1 Setup, Principal Strata, and Parameter of Interest

Assume we have a random sample of size \( n \) from a large population. For each unit \( i \) in the sample, let \( Z_i = z \in \{0, 1\} \) indicate whether the unit was randomly assigned to the treatment group \( (Z_i = 1) \) or to the control group \( (Z_i = 0) \), let \( D_i = d \in \{0, 1\} \) indicate whether individual \( i \) actually received the treatment \( (D_i = 1) \) or not \( (D_i = 0) \), and let \( S_i = s \in \{0, 1\} \) be a post-treatment sample selection indicator for whether the latent outcome variable \( Y_i \) is observed \( (S_i = 1) \) or not \( (S_i = 0) \). In addition, let \( D(z) \) denote the potential compliance behavior as a function of the treatment assignment, and let \( S(z, d) \) and \( Y^*(z, d) \) denote the potential values of the selection indicator and the latent outcome, respectively, as a function of the treatment assignment \( (z) \) and the treatment received \( (d) \). We observe \( \{Z_i, D_i(Z_i), S_i(Z_i, D_i(Z_i))\} \) for all units, whereas the observed outcome is \( Y_i = Y_i^*(Z_i, D_i(Z_i)) \) if \( S_i = 1 \) and is missing if \( S_i = 0 \). Our notation implicitly imposes the stable unit treatment value assumption (SUTVA) (Rubin 1978, 1980, 1990), which implies that there are no different versions of the treatment and there is no interference between individuals.

In our application, \( Z_i \) specifies whether individual \( i \) was randomly assigned to participate or not in JC, \( D_i \) denotes whether the individual actually enrolled in JC, \( S_i \) specifies whether individual \( i \) is employed or not, and \( Y_i^* \) is the offered market wage. There is disagreement in the literature on whether it is possible to define wages for individuals who are unemployed. For example, ZRM, Zhang et al. (2009), and Frumento et al. (2012) argue that wages are well-
defined only for those who are employed and, thus, do not consider parameters or assumptions involving wages for the unemployed (which are never observed). Here, we follow Blundell et al. (2007) and Lee (2009), and interpret $Y^*$ for unemployed individuals as the wages rejected by those who do not take up employment. However, our main results are unaffected by the way we treat $Y^*$ when $S = 0$ because our parameter of interest considers only individuals who would be employed under both treatment arms.

To simplify notation, in what follows we omit the subscript $i$ unless deemed necessary. We address noncompliance by using randomization into JC ($Z$) as an IV for JC enrollment ($D$). As in AIR, we impose the following assumptions:

**Assumption 1 (Randomly Assigned Instrument).** \{Y^*(z,d), S(z,d), D(z)\} is independent of $Z$ for all $z, d \in \{0,1\}$.

**Assumption 2 (Exclusion Restriction of $Z$).** $Y^*(z,d) = Y^*(z',d) = Y^*(d)$ and $S(z,d) = S(z',d) = S(d)$ for all $z, d \in \{0,1\}$.

**Assumption 3 (Nonzero Average Effect of $Z$ on $D$).** $E[D (1) - D (0)] \neq 0$.

Assumption 2 states that any effect of the IV ($Z$) on the potential outcomes $Y^*$ and on the potential sample selection indicator $S$ must be via an effect of $Z$ on the treatment $D$. In our application, it requires that randomization affects potential wages and employment only through its effect on JC enrollment. Assumption 2 allows us to write the potential variables $Y^*(z,d)$ and $S(z,d)$ as a function of the treatment $d$ only. Assumption 3 requires the IV (randomization) to have a non-zero average effect on the probability of actually receiving treatment (JC enrollment). Note that we require $Z$ to be a valid IV for both $Y^*$ and $S$.

We derive bounds for the effect of JC enrollment on wages accounting for sample selection and noncompliance within a principal stratification framework (Frangakis and Rubin, 2002). This framework is useful for studying causal effects when controlling for intermediate variables that have been affected by the treatment assignment ($Z$). The basic principal stratification with respect to $Z$ is a partition of the population into groups whose members share the same potential values of the intermediate variable under each value of $Z$. Since membership to a
particular stratum is not affected by treatment assignment, comparisons of potential outcomes within a given stratum yield causal effects.

The intermediate variables we want to control for are the compliance behavior \((D)\) and the sample-selection indicator \((S)\). In our setting, the principal strata are defined by the joint potential values of \(\{D(z = 0), D(z = 1)\} \times \{S(z = 0), S(z = 1)\}\). We define the following subpopulations: 
- \(a = \{i : D_i(0) = D_i(1) = 1\}\), the “always-takers”;
- \(n = \{i : D_i(0) = D_i(1) = 0\}\), the “never-takers”;
- \(c = \{i : D_i(0) = 0, D_i(1) = 1\}\), the “compliers”;
- \(d = \{i : D_i(0) = 1, D_i(1) = 0\}\), the “defiers”; as well as
- \(EE = \{i : S_i(0) = S_i(1) = 1\}\), the “always-employed”, those who would be employed regardless of treatment assignment;
- \(NN = \{i : S_i(0) = S_i(1) = 0\}\), the “never-employed”, those who would be unemployed regardless of treatment assignment for them;
- \(NE = \{i : S_i(0) = 0, S_i(1) = 1\}\), those who would be employed only if assigned to the treatment group;
- \(EN = \{i : S_i(0) = 1, S_i(1) = 0\}\), those who would be employed only if assigned to the control group.

In total, we have sixteen strata: \(\{a, n, c, d\} \times \{EE, NN, NE, EN\}\). These strata are similar to those in Frumento et al. (2012), and they result from combining the strata employed in AIR to account for noncompliance with those used in ZRM to address sample selection (see also Mattei and Mealli, 2007).

An important characteristic of principal strata is that they are latent subpopulations, meaning that, in general, we cannot observe to which stratum each individual belongs. Thus, additional assumptions are usually imposed to identify effects of interest by reducing the number of strata that exist in the population. Assumption 2 implies that the strata \(aNE, aEN, nNE,\) and \(nEN\) do not exist because, for these four strata, there exists an effect of the treatment assignment \((Z)\) on employment \((S)\) that is not through JC enrollment (since for them \(D(1) = D(0)\)), which contradicts the exclusion restriction of \(Z\). We also impose the following assumption, also used by AIR, which further reduces the number of existing strata.

**Assumption 4** (Individual-Level Monotonicity of \(D\) in \(Z\)). \(D_i(1) \geq D_i(0)\) for all \(i\).

Assumption 4 rules out the existence of defiers, thus eliminating the strata \(dEE, dNN, dEN,\) and \(dNE\). In our application, it assumes no individual would enroll in JC only if assigned to the control group. Assumption 4 implies that the average effect of \(Z\) on \(D\), which is
point identified, is non-negative. In the absence of sample selection, Imbens and Angrist (1994) and AIR show that under Assumptions 1 through 4 IV estimators point identify the average treatment effect of $D$ on $Y^*$ for the compliers. If sample selection is present, however, those assumptions are not enough to point identify an average effect of $D$ on $Y^*$.

In this paper, our parameter of interest (also considered in Frumento et al., 2012) is the average effect for the intersection of the subpopulation AIR focus on when accounting for noncompliance with the one ZRM and Lee (2009) focus on when addressing sample selection. More specifically, we focus on the average treatment effect of enrolling in JC ($D$) on latent wages ($Y^*$) for the “always-employed compliers” (i.e., the $cEE$ stratum):

$$\Delta = E[Y^*(1) - Y^*(0)|cEE].$$

The focus is on the effect for the $cEE$ stratum because, from the definition of the different strata, this is the only stratum for which wages are non-missing for individuals who enrolled and did not enroll in JC (after imposing Assumption 4). Hence, the effect for this stratum is the only one that is well-defined whether or not we are willing to define wages for unemployed individuals, and is also the only one for which bounds can be derived without imposing assumptions on wages that are never observed in our setting. From the perspective that wages are well-defined only for employed individuals (e.g., ZMR), a causal effect of JC on wages is well-defined only for $EE$ individuals (e.g., $NE$ individuals’ wages are missing under the control treatment and their effect is thus undefined). Even if one assumes wages are well-defined for the unemployed (as we do), wages are never observed in at least one treatment arm for non-$EE$ individuals. Moreover, among the $EE$ individuals, wages are observed under enrollment and non-enrollment in JC only for compliers (e.g., $nEE$ members’ wages under enrollment are never observed). Thus, bounding effects for strata other than the $cEE$ stratum requires assumptions about quantities that are never observed regardless of treatment assignment, and there is disagreement in the literature on whether such assumptions can be made. For example, the original principal stratification framework of Frangakis and Rubin (2002) only employs potential outcomes that can be observed in the data depending on the treatment assignment, while other authors (e.g., Flores and Flores-Lagunes 2010, 2013; Chen et al., 2012; Huber and Mellace, 2013) also use
a priori counterfactual outcomes that are never observed regardless of treatment assignment (e.g., the \( nEE \) members’ wages if we could “force” them to enroll in JC) and thus are “entirely hypothetical” (Rubin, 1990). Following the principal stratification literature, in the next two sections we avoid that type of assumptions to derive our bounds on (1). Lastly, note that the stratum \( cEE \) can also be interpreted as those compliers who would be always employed regardless of treatment receipt, since for compliers \( Z = D \).

### 2.2 Bounds under an Additional Monotonicity Assumption

We impose the following assumption.

**Assumption 5** *(Individual-Level Monotonicity of \( S \) in \( D \)). \( S_i(1) \geq S_i(0) \) for compliers.*

Assumption 5 states that there is a non-negative effect of \( D \) on \( S \) for every complier. It differs from the monotonicity assumption used in ZRM and Lee (2009) in that monotonicity is imposed in the actual treatment received (rather than in the treatment assigned) and only for compliers (rather than for all individuals). A testable implication of Assumption 5 is that the average effect of \( D \) on \( S \) for compliers, which is point identified under Assumptions 1 through 4, is non-negative. Individual-level monotonicity assumptions are common in the partial identification literature (e.g., Manski and Pepper, 2000; Flores and Flores-Lagunes, 2013). In our application, Assumption 5 states that there is a non-negative effect of JC on employment for every complier. While this assumption is weaker than that in ZRM and Lee (2009) (as it only applies to compliers), it is not innocuous because it is imposed at the individual level and thus it requires that no individual complier has a negative effect of JC on employment. We discuss potential threats to the plausibility of this assumption in Section 3.2.

From the strata remaining after imposing Assumptions 1 through 4, Assumption 5 rules out the existence of the \( cEN \) stratum. Hence, under these five assumptions there are seven strata in the population: \( aEE \), \( aNN \), \( nEE \), \( nNN \), \( cEE \), \( cNN \) and \( cNE \). The relationship between these seven strata and the observed groups defined by the values of \( \{Z, D(Z), S(Z, D(Z))\} \) is given in Table 1. This table illustrates that the \( cEE \) stratum is the only one for which wages are observed under both \( D = 0 \) and \( D = 1 \).
Table 1: Observed Groups and Principal Strata

<table>
<thead>
<tr>
<th>Z = 0</th>
<th>Z = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Let \( \pi_k \) denote the proportion of stratum \( k \) in the population, and let \( p_{ds|z} = \Pr(D = d, S = s|Z = z) \) and \( q_{s|z} = \Pr(S = s|Z = z) \). Under Assumptions 1 through 5, we can identify \( \pi_k \) for all strata as:

\[
\begin{align*}
\pi_{aNN} &= p_{10|0}, \\
\pi_{aEE} &= p_{11|0}, \\
\pi_{nNN} &= p_{00|1}, \\
\pi_{nEE} &= p_{01|1}, \\
\pi_{cEE} &= p_{01|0} - p_{00|1}, \\
\pi_{cNN} &= p_{10|1} - p_{10|0}, \\
\pi_{cNE} &= q_{1|0} - q_{1|0}.
\end{align*}
\]

Similarly, letting \( Y_{zds} = \mathbb{E}[Y|Z = z, D = d, S = s] \), we can write the mean outcomes for the observed cells with \( S = 1 \) as a function of mean potential outcomes for different strata as:

\[
Y_{101} = \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{nEE}} E[Y^*(0)|nEE] + \frac{\pi_{nEE}}{\pi_{cEE} + \pi_{nEE}}, \quad (2)
\]

\[
Y_{111} = \frac{\pi_{cEE} \cdot E[Y^*(1)|cEE]}{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}} + \frac{\pi_{cEE} \cdot E[Y^*(1)|cNE]}{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}} + \frac{\pi_{aEE} \cdot E[Y^*(1)|aEE]}{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}}, \quad (3)
\]

In addition to \( E[Y^*(1)|aEE] \) and \( E[Y^*(0)|nEE] \), \( E[Y^*(0)|cEE] \) is also identified, from (2), by:

\[
E[Y^*(0)|cEE] = \frac{p_{01|0}}{p_{01|0} - p_{00|1}} Y_{101} - \frac{p_{01|1}}{p_{01|0} - p_{00|1}} Y_{101}. \quad (4)
\]

Therefore, one of the terms of \( \Delta \) in (1) is point identified. However, the term \( E[Y^*(1)|cEE] \) is not point identified because two of the conditional means in (3) are not point identified. Next, we construct bounds for \( E[Y^*(1)|cEE] \) and, thus, \( \Delta \).

In the absence of noncompliance, ZRM and Lee (2009) construct bounds for the non-point identified mean of \( Y^*(1) \) for always-employed (EE) individuals from a cell containing only two strata. To illustrate the idea behind their trimming bounds, suppose there were no individuals in the \( aEE \) stratum (\( \pi_{aEE} = 0 \)), so that the cell \( \{Z = 1, D = 1, S = 1\} \) had only two strata, \( cEE \) and \( cNE \). Then, \( E[Y^*(1)|cEE] \) would be bounded from above (below) by the mean of \( Y \) for the fraction \( \pi_{cEE}/(\pi_{cEE} + \pi_{cNE}) \) of the largest (smallest) values of \( Y \) for those individuals.
in that cell. A key difference between the bounds derived in those papers and ours is that in our setting the bounds for $E[Y^*(1)|cEE]$ are derived from a cell containing three strata.

Although $\bar{Y}^{111}$ is a weighted average of the mean of $Y^*(1)$ for three strata (equation 3), $E[Y^*(1)|aEE]$ is point identified. To motivate how we construct our bounds, we can think of the problem as finding “worst-case” scenarios for $E[Y^*(1)|cEE]$ subject to the constraint that $\bar{Y}^{011} = E[Y^*(1)|aEE]$. To bound $E[Y^*(1)|cEE]$, we solve the unconstrained problem first and check whether the value of $E[Y^*(1)|aEE]$ implied by this solution can satisfy the constraint, in which case the unconstrained solution is the solution to the constrained problem. If not, we impose the constraint first and then obtain the solution to the constrained problem.

To describe our bounds, let $y^{111}_r$ be the $r$-th quantile of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$, and let $\bar{Y}(y^{111}_r \leq Y \leq y^{111}_r) \equiv E[Y|Z = 1, D = 1, S = 1, y^{111}_r \leq Y \leq y^{111}_r]$ be the mean of the outcomes between the $r$-th and $r$-th quantiles of $Y$ in that cell. Suppose we want to derive the lower bound for $E[Y^*(1)|cEE]$. We first consider the problem without the constraint. In this case, we can apply the trimming procedure in ZRM and Lee (2009) and bound $E[Y^*(1)|cEE]$ from below by the expected value of $Y$ for the $\pi_{cEE}/p_{11|1}$ fraction of the smallest values of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$, or, $\bar{Y}(Y \leq y^{111}_{\pi_{cEE}/p_{11|1}})$, where $p_{11|1} = \pi_{cEE}+\pi_{cNE}+\pi_{aEE}$. Next, we check whether this solution is consistent with the constraint $\bar{Y}^{011} = E[Y^*(1)|aEE]$. To do this, we construct the “worst-case” scenario lower bound for $E[Y^*(1)|aEE]$, call it $\bar{y}_{aEE}$, implied by the unconstrained solution by placing all the $aEE$ individuals at the bottom of the remaining observations in the cell $\{Z = 1, D = 1, S = 1\}$. This yields $\bar{y}_{aEE} = \bar{Y}(y^{111}_{\pi_{cEE}/p_{11|1}} \leq Y \leq y^{111}_{1-(\pi_{cNE}/p_{11|1})})$. If $\bar{y}_{aEE} \leq \bar{Y}^{011}$, the unconstrained solution is consistent with the constraint and the lower bound for $E[Y^*(1)|cEE]$ is $\bar{Y}(Y \leq y^{111}_{\pi_{cEE}/p_{11|1}})$; otherwise, the unconstrained solution is inconsistent with $\bar{Y}^{011} = E[Y^*(1)|aEE]$. Intuitively, having $\bar{y}_{aEE} > \bar{Y}^{011}$ implies that some observations from the $aEE$ stratum must be at the bottom $\pi_{cEE}/p_{11|1}$ fraction of the smallest values of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$ and thus we can improve upon the lower bound $\bar{Y}(Y \leq y^{111}_{\pi_{cEE}/p_{11|1}})$ for $E[Y^*(1)|cEE]$. In this case, we construct the “worst-case” scenario lower bound for $E[Y^*(1)|cEE]$ by placing all the observations in the $aEE$ and $cEE$ strata at the bottom of the distribution of $Y$ in that cell. Thus, if $\bar{y}_{aEE} > \bar{Y}^{011}$, the lower
bound for $E[Y^*(1)|cEE]$, call it $LY_{1,cEE}$, is derived from the equation:

$$
\overline{Y}(Y \leq y_{111}^{111 - (\pi_{cEE}/p_{111})}) = \frac{\pi_{cEE}}{\pi_{cEE} + \pi_{aEE}} LY_{1,cEE} + \frac{\pi_{aEE}}{\pi_{cEE} + \pi_{aEE}} \overline{Y}^{011},
$$

(5)

where $\overline{Y}(Y \leq y_{111}^{111 - (\pi_{cEE}/p_{111})})$ is the mean of $Y$ for the $1 - (\pi_{cEE}/p_{111})$ fraction of the smallest values of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$.

Note that if $\overline{Y}_{aEE} \leq \overline{Y}^{011}$, the lower bound derived from (5) will be smaller than the lower bound derived without using the information on $E[Y^*(1)|aEE]$, $\overline{Y}(Y \leq y_{111}^{111 - (\pi_{cEE}/p_{111})})$. Intuitively, if $\overline{Y}^{011} = E[Y^*(1)|aEE]$ is very large, it is impossible that all $aEE$ individuals are at the bottom $1 - (\pi_{cEE}/p_{111})$ fraction of the smallest values of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$, and the lower bound derived from (5) will be lower than $\overline{Y}(Y \leq y_{111}^{111 - (\pi_{cEE}/p_{111})})$. In general, we can write the lower bound for $E[Y^*(1)|cEE]$ as the maximum of those two lower bounds.

The upper bound for $E[Y^*(1)|cEE]$ is derived in a similar way by placing the observations in the corresponding strata in the upper part of the distribution of $Y$ in the cell $\{Z = 1, D = 1, S = 1\}$. The bounds for $E[Y^*(1)|cEE]$ are combined with the point identification of $E[Y^*(0)|cEE]$ in (4) to construct bounds for the average effect of the always-employed compliers, $\Delta$ in (1). Proposition 1 presents bounds for $\Delta$ under Assumptions 1 through 5.

**Proposition 1** If Assumptions 1 through 5 hold, then $L_{cEE} \leq \Delta \leq U_{cEE}$, where

$$
L_{cEE} = LY_{1,cEE} - \overline{Y}^{001} \frac{p_{01|0}}{p_{01|0} - p_{01|1}} + \overline{Y}^{101} \frac{p_{01|1}}{p_{01|0} - p_{01|1}},
$$

$$
U_{cEE} = UY_{1,cEE} - \overline{Y}^{001} \frac{p_{01|0}}{p_{01|0} - p_{01|1}} + \overline{Y}^{101} \frac{p_{01|1}}{p_{01|0} - p_{01|1}},
$$

$$
LY_{1,cEE} = \max \left\{ \overline{Y}(Y \leq y_{111}^{111 - \alpha_{cEE}}), \overline{Y}(Y \leq y_{111}^{111 - \alpha_{cEE}}) q_{1|0} - p_{01|1} p_{01|0} - p_{01|1} - \overline{Y}^{011} \frac{p_{11|0}}{p_{01|0} - p_{01|1}} \right\},
$$

$$
UY_{1,cEE} = \min \left\{ \overline{Y}(Y \geq y_{111}^{111 - \alpha_{cEE}}), \overline{Y}(Y \geq y_{111}^{111 - \alpha_{cEE}}) q_{1|0} - p_{01|1} p_{01|0} - p_{01|1} - \overline{Y}^{011} \frac{p_{11|0}}{p_{01|0} - p_{01|1}} \right\},
$$

$$
\alpha_{cEE} = \frac{\pi_{cEE}}{p_{11|1}} = \frac{p_{01|0} - p_{01|1}}{p_{11|1}}, \text{ and } \alpha_{cEE} = \frac{\pi_{cEE}}{p_{11|1}} = \frac{q_{1|1} - q_{1|0}}{p_{11|1}}.
$$

**Proof.** See Internet Appendix.
2.3 Bounds under Mean Dominance

We now consider an assumption that narrows the bounds presented in Proposition 1.

**Assumption 6 (Mean Dominance).** \( E[Y^*(1)|cEE] \geq E[Y^*(1)|cNE] \).

Assumption 6 comes from the notion that some strata are likely to have more favorable characteristics and thus better potential outcomes than others. In our application, this assumption states that the mean potential wage under treatment of the always-employed compliers \((cEE)\) stratum) is greater than or equal to that of those compliers who would be employed only if they enrolled in JC \((cNE)\) stratum). Assumption 6 implies a positive correlation between employment and wages, which is supported by standard models of labor supply. ZRM and Huber and Mellace (2013) consider stochastic-dominance versions of Assumption 6 that require weakly dominance at any rank of the potential outcome distribution, instead of only at the mean. Stochastic dominance is much stronger than needed for our purposes.

In practice, the bounds under both Assumptions 5 and 6 may be preferred over those under only Assumption 5 because they are narrower, in which case it is important to have a sense about the plausibility of Assumption 6. Although it is not directly testable, it is possible to get indirect evidence about its plausibility by comparing the average baseline characteristics of the \(cEE\) and \(cNE\) strata that are closely related to the outcome of interest (e.g., pre-randomization outcome values). Assumption 6 is less likely to hold if these comparisons suggest that the \(cNE\) stratum has better characteristics at the baseline than does the \(cEE\) stratum. Under Assumptions 1 through 5 the average baseline characteristics of all the strata are point identified from the observed average baseline characteristics of the observed groups \(\{Z, D, S\}\) in Table 1, as each of these eight observed means is a weighted average of the mean baseline characteristics of the seven strata (see, for reference, equations (2) and (3)), with the weights being also point identified. Since the number of moment conditions implied by these equations is greater than the number of parameters, in our application we use generalized method of moments (GMM) to estimate the mean baseline characteristics of all seven strata. GMM minimizes a squared Euclidean distance of the sample analogues of the moment conditions to their population value of zero (e.g., Newey and McFadden, 1994). We provide further details in the Internet Appendix.
Assumption 6 tightens the bounds in Proposition 1 by increasing the lower bound on $E[Y^{*}(1)|cEE]$. Similar to equation (3), we can write

$$Y^{111} = \frac{(\pi_{cEE} + \pi_{cNE}) \cdot E[Y^{*}(1)|cEE, cNE]}{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}} + \frac{\pi_{aEE} \cdot E[Y^{*}(1)|aEE]}{\pi_{cEE} + \pi_{cNE} + \pi_{aEE}},$$

(6)

where the stratum proportions and $E[Y^{*}(1)|aEE]$ are point identified. Assumption 6 implies $E[Y^{*}(1)|cEE] \geq E[Y^{*}(1)|cEE, cNE]$, which provides a lower bound for $E[Y^{*}(1)|cEE]$ that is greater than or equal to the one in Proposition 1. Proposition 2 presents bounds for $\Delta$ under Assumptions 1 through 6.

**Proposition 2** If Assumptions 1 through 6 hold, then $\overline{L}_{cEE} \leq \Delta \leq U_{cEE}$, where $U_{cEE}$ is as given in Proposition 1 and $\overline{L}_{cEE}$ equals:

$$\overline{L}_{cEE} = \overline{LY}_{1,cEE} - \overline{Y}^{001} \frac{p_{01|0}}{p_{01|0} - p_{01|1}} + \overline{Y}^{101} \frac{p_{01|1}}{p_{01|0} - p_{01|1}},$$

with

$$\overline{LY}_{1,cEE} = \frac{p_{11|1}Y^{111} - p_{11|0}Y^{011}}{p_{11|1} - p_{11|0}}.$$

**Proof.** See Internet Appendix.

### 2.4 Remarks

**Remark 1.** It is possible to construct bounds on $\Delta$ without Assumption 5, in which case the stratum $cEN$ is not ruled out and appears in the observed cells $\{Z = 0, D = 0, S = 1\}$ and $\{Z = 1, D = 1, S = 0\}$ in Table 1. Although the proportions of the strata $aEE$, $aNN$, $nEE$, and $nNN$ are still point identified, neither the proportions of the strata $cEE$, $cNN$, $cNE$, and $cEN$ nor the term $E[Y^{*}(0)|cEE]$ is now point identified. To construct bounds for $\Delta$ in this case, we can combine our approach in Section 2.2 with that followed by Zhang and Rubin (2003), ZRM, Imai (2008), and Huber and Mellace (2013) in a setting without noncompliance. The main idea for constructing their bounds on the average effect for the $EE$ stratum is to consider “worst-case” scenarios that are consistent with the possible values $\pi_{EE}$ can take based on the data, with the bounds being obtained at the minimal value $\pi_{EE}$ can take (Huber and Mellace, 2013). In our setting, the worst-case scenarios for $\Delta$ occur when $\pi_{cEE}$ is at its minimum value that is consistent with the data. The minimum possible value of $\pi_{cEE}$, calculated from the cells
in Table 1, is \( \pi_{cEE} \geq \max(0, p_{01|0} - p_{01|1} - p_{10|1} + p_{10|0}) \). As in Lee (2009), the bounds on \( \Delta \) are well-defined only if \( \pi_{cEE} > 0 \), which implies \( \pi_{cEE} \) is minimized at \( p_{01|0} - p_{01|1} - p_{10|1} + p_{10|0} \). Given this lower bound for \( \pi_{cEE} \), the same approach as in Section 2.2 can be used to derive bounds on \( \Delta \) by constructing bounds for \( E[Y^*(1)|cEE] \) and \( E[Y^*(0)|cEE] \). However, these bounds will be wider than those in Proposition 1 and may be uninformative in practice, as illustrated by Blanco et al. (2013) in a setting without noncompliance.

**Remark 2.** In the absence of Assumptions 5 and 6, the lower bound for \( \Delta \) in Proposition 2, \( L_{cEE} \), provides information for another parameter of interest: \( \beta = E[Y^*(1) - Y^*(0)|cEE, cNE] \), the average effect of \( D \) on \( Y^* \) for the \( cEE \) and \( cNE \) strata. \( L_{cEE} \) gives a lower bound for \( \beta \) under Assumptions 1 through 4 and the following assumption:

**Assumption 5’** \( E[Y^*(0)|cEE, cEN] \geq E[Y^*(0)|cEE, cNE] \).

This assumption states that the mean of \( Y^*(0) \) (i.e., the latent wage under no enrollment in JC) for compliers who would be employed if they did not attend JC (\( cEE \) and \( cEN \)) is greater than or equal to that for compliers who would be employed if they did attend JC (\( cEE \) and \( cNE \)). This assumption exploits the positive correlation between employment and wages implied by standard models of labor supply, as the \( cEN \) are employed under the control treatment but the \( cNE \) are not. Note that in this case one needs to be willing to define wages for the unemployed and make assumptions about never-observed counterfactual wages.

To see why \( L_{cEE} \) provides a lower bound for \( \beta \), note that \( E[Y^*(1)|cEE, cNE] \) is point identified from (6) and equals \( YY_{1,cEE} \) (defined in Proposition 2), while Assumption 5’ implies that an upper bound for \( E[Y^*(0)|cEE, cNE] \) is given by \( E[Y^*(0)|cEE, cEN] \), which is point identified as \( E[Y^*(0)|cEE, cEN] = (Y_{001} - Y_{101}^t) / (p_{01|0} - p_{01|1}) \) from

\[
Y_{001}^{001} = \frac{(\pi_{cEE} + \pi_{cEN}) \cdot E[Y^*(0)|cEE, cEN]}{\pi_{cEE} + \pi_{cEN} + \pi_{nEE}} + \frac{\pi_{nEE} \cdot E[Y^*(0)|nEE]}{\pi_{cEE} + \pi_{cEN} + \pi_{nEE}} \tag{7}
\]

In practice, if selected compliers under control (\( cEE \) and \( cEN \) strata) have more favorable characteristics than selected compliers under treatment (\( cEE \) and \( cNE \) strata), we would expect the former group to be more likely to experience better outcomes than the latter in the absence of treatment. Thus, indirect evidence about the plausibility of Assumption 5’ can
be obtained by comparing $E[X|cEE, cEN]$ and $E[X|cEE, cNE]$, where $X$ denotes relevant baseline characteristics (e.g., pre-randomization outcomes). Similarly, it is important to have a sense of the type of individuals who may belong to the $cEN$ stratum. In our application these could be compliers who, for example, if they enrolled in JC would pursue an associate or other academic degree after leaving JC and would appear as unemployed at the time the outcome is measured, while, if not enrolled in JC, they may have no other choice but to work to cover their basic needs (JC also provides room, board, and a stipend) and are able to find a job.

In sum, in applications where Assumption 5 is difficult to justify, Assumption 5’ may be an attractive alternative. Moreover, the mixture of the strata $cEE$ and $cNE$ seems to be an interesting subpopulation, since these are individuals who would comply with the treatment assignment and be employed if they attended JC.

### 2.5 Estimation and Inference

The bounds derived in Sections 2.2 and 2.3 contain minimum (min) and maximum (max) operators, which create problems for standard estimation and inference procedures. Sample analog estimates of those bounds tend to be tighter than the true bounds due to the concavity (convexity) of the min (max) function, and the asymptotic distribution of the bound estimators is usually unavailable. Hirano and Porter (2012) show that for parameters that are nonsmooth functionals of the underlying data distribution, such as those having min or max operators, there exist no locally asymptotically unbiased estimators and no regular estimators. A growing literature has focused on developing valid estimation and inference procedures for this type of bounds (e.g., Tamer, 2010). We use the methodology proposed by Chernozhukov, Lee, and Rosen (2013) (hereafter CLR) to obtain confidence regions for the true parameter value and half-median unbiased estimators for the lower and upper bounds. The half-median-unbiasedness property means that the upper bound estimator exceeds the true value of the upper bound with probability at least one half asymptotically, while the reverse holds for the lower bound.

To briefly describe CLR’s procedure, let the bounds for a parameter $\theta_0$ (e.g., $\Delta$) be given by $[\theta_0^l, \theta_0^u]$, where $\theta_0^l = \max_{v \in V} \{1, \ldots, m\} \theta^l(v)$ and $\theta_0^u = \min_{v \in V} \{1, \ldots, m\} \theta^u(v)$. CLR refer to $\theta^l(v)$ and $\theta^u(v)$ as bounding functions. In our setting, $v$ indexes the bounding functions, while
and $m^u$ give, respectively, the number of terms inside the max and min operators. For example, the upper bound for $\theta_0 = E[Y^*(1) | c_{E,E}]$ in Proposition 1, $UY_{1,E,E}$, can be written as 

$$\theta_0^u = \min_{v \in V^u = \{1, 2\}} \theta^u(v),$$

with $\theta^u(1) = \bar{Y}(Y \geq y_{1,\alpha_{c,E,E}}^{11})$ and $\theta^u(2) = \bar{Y}(Y \geq y_{1,\alpha_{c,E,E}}^{11} - \frac{Y_0^{11} - p_{11}^{10}}{p_{01}^{10} - p_{01}^{11}}.$ In our case, sample analog estimators of the bounding functions are known to be consistent and asymptotically normally distributed, as they are simple functions of conditional probabilities, means, and trimmed means (Lee, 2009; Newey and McFadden, 1994).

CLR address the issues related to estimation and inference for the bounds $[\theta_0^l, \theta_0^u]$ by employing precision-corrected estimates of the bounding functions before applying the min and max operators. The precision adjustment consists of adding to each estimated bounding function its pointwise standard error times an appropriate critical value, $\kappa(p)$, so that estimates with higher standard errors receive larger adjustments. Depending on the choice of $\kappa(p)$, it is possible to obtain confidence regions for either the identified set or the true parameter value, and half-median unbiased estimators for the lower and upper bounds.

More specifically, the precision-corrected estimator of the upper bound $\theta_0^u$ is given by

$$\hat{\theta}^u(p) = \min_{v \in V^u} \left[ \hat{\theta}^u(v) + \kappa_{\hat{V}_n^u}(p)s^u(v) \right],$$

where $\hat{\theta}^u(v)$ is the sample analog estimator of $\theta^u(v)$ and $s^u(v)$ is its standard error. CLR compute the critical value $\kappa_{\hat{V}_n^u}(p)$ based on simulation methods and a preliminary estimator $\hat{V}_n^u$ of $V^u = \arg\min_{v \in V^u} \theta^u(v)$. Intuitively, $\hat{V}_n^u$ selects those bounding functions that are close enough to binding to affect the asymptotic distribution of the estimator of the upper bound. A precision-corrected estimator of the lower bound $\theta_0^l$ is obtained in a similar way. Further details on the CLR procedure and our implementation steps are provided in the Internet Appendix.

3 The Wage Effects of Job Corps

3.1 Data and Preliminary Analysis

We employ data from the National Job Corps Study (NJCS), a randomized experiment funded by the US Department of Labor to evaluate the effectiveness of JC. Eligible individuals who applied to JC for the first time between November 1994 and December 1995 (80,833 individuals) were randomly assigned into a treatment group and a control group. Individuals in the control
group (5,977) were embargoed from the program for three years, while those in the treatment group (74,856) were allowed to enroll in JC. The research sample (15,386), which consisted of all control individuals plus 9,409 randomly-selected individuals from the treatment group, was interviewed at random assignment and at 12, 30, and 48 months after random assignment.

The specific sample we employ is based on the one used by Lee (2009). This sample contains only individuals with non-missing values for weekly earnings and weekly hours worked for every week after randomization (9,145 individuals). We add to our data set enrollment information at week 208 (i.e., 48 months) after randomization. This is a binary indicator for whether or not the individual was ever enrolled in JC during the 208 weeks after random assignment. We drop 55 observations from Lee’s sample due to the missing enrollment variable, resulting in our final sample of 9,090 individuals: 3,599 in the control group and 5,491 in the treatment group. Wages at week 208 are obtained by dividing weekly earnings by weekly hours worked at that week, where wages being missing is equivalent to weekly hours worked being equal to zero. We regard an individual as unemployed when the wage is missing, and regard her as employed otherwise.

We use the NJCS design weights throughout our analysis because some subpopulations were randomized with different, but known, probabilities (Schochet et al., 2001).

Our definition of enrollment differs from that in Frumento et al. (2012), who consider enrollment within the first six months after randomization. As a result, the specific effect in (1) we estimate in this application differs from theirs because we estimate the effect for the \( cEE \) stratum of ever participating in JC within the 208 weeks (rather than within six months) after being offered to enroll in JC relative to never enrolling in JC within the same period. Moreover, the complier subpopulations also differ because in our case the compliers are those who would enroll in JC within the 208 weeks (rather than within six months) after randomization only if assigned to enroll. We define enrollment as above to take into consideration late noncompliance behavior and to illustrate the results discussed in the previous sections, since in the setting of Frumento et al. (2012) our bounds reduce to those in ZRM and Lee (2009) (they rule out the existence of always takers by dropping from their sample control individuals who enrolled in JC within the first six months after randomization –about 1% of controls– and by taking controls who enrolled in JC afterwards as control group members). Our enrollment definition
also differs from that in the NJCS reports by Schochet et al. (2001) and Burghardt et al. (2001), who consider enrollment within the three-year embargo period. Importantly, because of the relatively small fraction of control individuals who enroll in JC (in our sample, 4.4%) and the fact that most individuals assigned to enroll in JC do so within six months (96.23%), the bound estimates shown below remain almost the same if we define enrollment as in the NJCS reports (in which case there are always takers) or as in Frumento et al. (2012). The results for these two cases are provided in the Internet Appendix.

In addition, note that the end of the embargo in year three could potentially create issues with the way we define enrollment because, while 1.2 percent of controls enroll in JC before the end of the embargo, 3.2 percent enroll afterwards. This late enrollment by some of the controls could affect our estimated bounds. For example, relative to controls who enrolled before the embargo and had been out of JC for a longer period, late enrolled controls may be less likely to be employed and their wages are likely to be lower at week 208, which could affect our estimates of the stratum proportions (e.g., $\pi_{aE}$) and $E[Y(1)|aE]$, respectively. While it is difficult to precisely determine the way our estimated bounds would be affected, the fact that our estimated bounds are robust to the enrollment definition used suggests that this potential issue is unlikely to greatly affect the results shown below.

The first four columns of Table 2 report the averages of selected baseline characteristics for our entire sample by treatment assignment status. The averages of the complete set of baseline characteristics are presented in the Internet Appendix. The pre-treatment variables include demographic characteristics, education and background variables, and employment and earnings information at the baseline and in the year prior to randomization. As expected, the average pre-treatment characteristics of the treatment and control groups are very similar (the difference in means is statistically different from zero at the five percent level for only one—weekly hours worked at baseline—of the 31 pre-treatment variables shown in the Internet Appendix). Thus, our sample maintains the balance of baseline variables between both groups.

The first three columns of Table 3 show the averages of relevant post-treatment variables based on treatment assignment status, along with their differences, at week 208 after randomization. The first row shows that by week 208, 73.8 percent of those assigned to the treatment
Table 2: Summary Statistics of Selected Baseline Variables

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>Non-Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z=1</td>
<td>Z=0</td>
</tr>
<tr>
<td>Female</td>
<td>.454</td>
<td>.458</td>
</tr>
<tr>
<td>Age at Baseline</td>
<td>18.44</td>
<td>18.35</td>
</tr>
<tr>
<td>White, Non-hispanic</td>
<td>.265</td>
<td>.263</td>
</tr>
<tr>
<td>Black, Non-Hispanic</td>
<td>.494</td>
<td>.491</td>
</tr>
<tr>
<td>Hispanic</td>
<td>.169</td>
<td>.172</td>
</tr>
<tr>
<td>Never married</td>
<td>.917</td>
<td>.916</td>
</tr>
<tr>
<td>Has Children</td>
<td>.189</td>
<td>.193</td>
</tr>
<tr>
<td>Education (Years of Schooling)</td>
<td>10.12</td>
<td>10.11</td>
</tr>
<tr>
<td>Ever Arrested</td>
<td>.248</td>
<td>.249</td>
</tr>
<tr>
<td>Personal Income: &lt;3000</td>
<td>.788</td>
<td>.789</td>
</tr>
</tbody>
</table>

**At Baseline:**

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>Non-Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have job</td>
<td>.198</td>
<td>.192</td>
</tr>
<tr>
<td>Had job, Prev. Yr.</td>
<td>.635</td>
<td>.627</td>
</tr>
<tr>
<td>Months employed, Prev. Yr.</td>
<td>3.603</td>
<td>3.530</td>
</tr>
<tr>
<td>Earnings, Prev. Yr.</td>
<td>2911.0</td>
<td>2810.5</td>
</tr>
</tbody>
</table>

Numbers of observations: 5491 3599 Total: 9090 4551 2978 Total: 7529

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.

group and 4.4 percent of the control group had ever enrolled in JC. The difference in these two numbers, which equals the proportion of compliers, is 69.4 percent. The rest of the rows in Table 3 present the intention-to-treat (ITT) effect and the (local) average treatment effect for compliers (LATE) on various labor market outcomes at week 208. The second type of effects adjust for noncompliance by employing the randomly assigned treatment as an instrument for JC enrollment. The ITT and LATE effects of JC on weekly hours worked, weekly earnings, and employment are positive and statistically significant, with the LATE estimates being larger than the ITT estimates by about 44, 44 and 50 percent, respectively. The estimated average effects of JC on earnings and employment for compliers is 39.9 dollars and 6 percentage points, respectively. These results are consistent with the findings in the NJCS (Burghardt et al., 2001). For reference, Table 3 also shows the estimated ITT and LATE effects of JC on ln(wage) for employed individuals. The LATE estimate implies an average effect of JC on ln(wage) for compliers of 0.053. However, these estimates are biased because of sample selection.
Table 3: Summary Statistics of Post-treatment Variables and Preliminary Effects

| Enrollment Variable: | Entire Sample | | | Non-Hispanics | | |
|----------------------|--------------|---|----------------|------------|---|----------------|------------|
| Z=1 | Z=0 | Diff.(Std.Err.) | Z=1 | Z=0 | Diff.(Std.Err.) |
| Ever enrolled in JC | .738 | .044 | .694**(.007) | .737 | .047 | .689**(.008) |
| **Intention-to-Treat (ITT) Effects:** | | | | | | |
| Hours per week | 27.80 | 25.83 | 1.967**(.559) | 28.05 | 25.53 | 2.523**(.617) |
| Earnings per week | 228.19 | 200.50 | 27.69**(.512) | 230.22 | 194.66 | 35.57**(5.555) |
| Employed | .607 | .566 | .041**(.011) | .609 | .559 | .050**(.012) |
| ln(wage) | 2.029 | 1.991 | .038**(.011) | 2.028 | 1.977 | .050**(.013) |

**Local Average Effect for Compliers, LATE (IV estimates):**

| | Entire Sample | | | Non-Hispanics | | |
|----------------------|--------------|---|----------------|------------|---|----------------|------------|
| Z=1 | Z=0 | Diff.(Std.Err.) | Z=1 | Z=0 | Diff.(Std.Err.) |
| Hours per week | 2.834**(.787) | 3.661**(.877) |
| Earnings per week | 39.90**(7.293) | 51.62**(8.007) |
| Employed | .060**(.015) | .072**(.017) |
| ln(wage) | .053**(.016) | .071**(0.017) |

Note: Outcomes are measured at Week 208 after randomization. Z is an indicator for whether the individual was randomly assigned to participate or not in JC. ** denotes that difference is statistically different from 0 at 5% level. Computations use design weights.

### 3.2 Assessment of Assumptions

Assumption 1 holds by design, and the results in Table 3 suggest that Assumption 3 is satisfied. Assumption 2 states that JC affects wages and employment only through its effect on JC enrollment. This assumption seems plausible and has been used before in the JC literature (e.g., Schochet, 2001; Schochet et al., 2001; Burghardt et al., 2001; Frumento et al., 2012). However, there could be threats to its validity. For example, the offer of a JC slot could affect the job search behavior of some individuals by increasing their reservation wages relative to what they would have been without the option to enroll in JC. Although this type of responses could affect the labor market outcomes in the short run, it seems reasonable to assume that treatment assignment has a negligible effect on the long run outcomes we consider through channels other than JC enrollment (see Schochet et al., 2001, and Frumento et al., 2012 for further discussion on the plausibility of this assumption in JC). Assumption 4 (no defiers) has also been used in the JC literature and appears plausible, as it seems highly unlikely that an individual would enroll in JC only if they were denied access to it.

Assumption 5 states that there is a non-negative effect of JC on employment for every individual complier at week 208. This assumption seems plausible given the objectives and
services provided by JC (e.g., academic and vocational training, job search assistance). Its testable implication that the LATE of JC on employment for compliers is non-negative is soundly supported by the data, as shown in Table 3. Nevertheless, there are two potential threats to the validity of Assumption 5. First, individuals who enroll in JC may be less likely to be employed while undergoing training than those who do not enroll (usually called the “lock-in” effect; van Ours, 2004). Second, trained individuals may raise their reservation wages because of the JC training, which may lead them to reject some job offers that they would otherwise accept if they had not received training. Both potential threats, however, are likely to become less relevant in the long run, as trained individuals are no longer “locked-in” away from employment, and individuals who chose to remain unemployed in the short run because of raising their reservation wages find jobs in the long run. Consistent with this view, Schochet et al. (2001) and Lee (2009) find negative effects of JC on employment in the short run, and positive effects in the long run. Thus, we focus our analysis on wages at week 208 after randomization, which is the latest wage measure available in the NJCS.

The likelihood-based analysis by Frumento et al. (2012) suggests that there may be a positive proportion of compliers in the population for whom JC has a negative effect on employment at week 208, even though this proportion falls over time after randomization. Hence, to further increase the plausibility of Assumption 5 in our application, we also consider a sample that excludes Hispanics. Hispanics were the only demographic group in the NJCS for which negative but statistically insignificant effects of JC on employment and earnings were found (Schochet et al., 2001), and therefore, Assumption 5 may not be appropriate for them. Schochet et al. (2001) find that the different results for Hispanics (relative to other groups) could not be explained by differences in features such as baseline characteristics, degree attainment, enrollment duration, or characteristics of the centers attended, whereas Flores-Lagunes et al. (2010) find that the lack of effect for Hispanics can be partly attributed to the higher local unemployment rates they face, and to the greater negative impact they experience from these unemployment rates. The last set of columns in Tables 2 and 3 present summary statistics and preliminary effects for the Non-Hispanics sample (7,529 individuals). As expected, the estimated ITT and LATE effects of JC on labor outcomes for Non-Hispanics are stronger than those for the entire sample.
Assumption 6 states that the mean potential outcome under treatment of the always-employed compliers (cEE stratum) is greater than or equal to that of individuals who would be employed only if they enrolled in JC (cNE stratum). As discussed in Section 2.3, to shed light on the plausibility of this assumption we compare average baseline characteristics of the cEE and cNE strata that are likely to be highly correlated to wages at week 208 for the entire and Non-Hispanics samples (the results are provided in the Internet Appendix). Focusing on the Non-Hispanics sample we find that, relative to individuals in the cNE stratum, individuals in the cEE stratum are more likely to be male and white, to have never been arrested at the baseline, and to have better labor market outcomes the year before randomization. These differences, however, are not statistically different from zero. We conclude from these results that the data do not provide indirect evidence against Assumption 6, and that the point estimates of the differences suggest that the assumption is plausible.

In sum, while not innocuous, Assumptions 1 through 6 appear reasonable in our application.

3.3 Empirical Results

**Bounds on ITT effect.** We start this section by bounding the average effect of being allowed to enroll in JC on wages (ITT effect) for the individuals who would always be employed regardless of treatment assignment, and then we bound the average effect of JC enrollment on wages for those compliers who would always be employed regardless of treatment receipt ($\Delta$ in (1)). The first parameter, which ignores noncompliance, is the one considered in ZRM and Lee (2009). In their setting, the principal strata are EE, NN, NE and EN, where the last stratum is ruled out by assuming monotonicity of $S$ in $Z$.

Table 4 presents estimation results for the average ITT effect of JC on ln(wage) for always-employed individuals (EE stratum). The first column of Table 4 presents results for our entire sample. The estimated proportion of the EE stratum in the population is 56.6 percent. Under the monotonicity of $S$ in $Z$ assumption, the estimated lower and upper bounds for the ITT effect of JC on ln(wage) for the EE stratum are $-0.022$ and $0.100$, respectively. These results are very similar to those obtained by Lee (2009) (they are not exactly the same because he uses a transformation of the wages –see footnote 13 in that paper– and we drop 55 observations
Table 4: Bounds for the ITT Effect of JC on ln(wage) for the EE Stratum

<table>
<thead>
<tr>
<th>Stratum Proportions:</th>
<th>Entire Sample</th>
<th>Non-Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always-employed (EE)</td>
<td>.566** (.009)</td>
<td>.559** (.009)</td>
</tr>
<tr>
<td>Never-employed (NN)</td>
<td>.393** (.007)</td>
<td>.391** (.007)</td>
</tr>
<tr>
<td>Employed only if assigned to program (NE)</td>
<td>.041** (.011)</td>
<td>.050** (.012)</td>
</tr>
<tr>
<td>(E[Y^*(Z = 0)</td>
<td>EE])</td>
<td>1.991** (.009)</td>
</tr>
<tr>
<td>Proportion of EE in cell (Z = 1, S = 1)</td>
<td>.932** (.017)</td>
<td>.918** (.019)</td>
</tr>
</tbody>
</table>

**Bounds under Monotonicity:**
- Lower bound for the ITT effect for EE stratum: -.022 (.016) - .018 (.017)
- Upper bound for the ITT effect for EE stratum: .100** (.014) - .119** (.015)
- Imbens and Manski 95% confidence interval: [-.048, .123] [-.047, .144]

**Bounds under Monotonicity and Mean Dominance:**
- Lower bound for the effect for EE stratum: .038** (.012) - .050** (.013)
- Upper bound for the ITT effect for EE stratum: .100** (.014) - .119** (.015)
- Imbens and Manski 95% confidence interval: [.019, .123] [.029, .144]

Note: Outcome is measured at week 208 after randomization. Numbers in parentheses are standard errors. ** denotes that estimate is statistically different from 0 at 5% level. Computations use design weights. Standard errors are calculated by a 5,000-repetition bootstrap. Imbens and Manski (2004) confidence intervals contain the true value of the parameter with a given probability.

from his sample due to missing enrollment information). As noted by Lee (2009), although the bounds include zero, they rule out plausible negative effects. Moreover, these lower bounds are implicitly based on the extreme and unintuitive assumption that wages are perfectly negatively correlated with the probability of being employed. This is contradicted by standard models of labor supply, in which individuals with higher wages are also more likely to be employed.

The last part of Table 4 presents bounds on the ITT effect when adding the mean dominance assumption that the mean potential wage under \(z = 1\) of the EE stratum is greater than or equal to that of the NE stratum. This assumption can be seen as a way to rule out the implausible extreme case mentioned above by implying a positive correlation between wages and employment. In this case, the estimated lower bound on the ITT effect of JC on ln(wage) for the EE stratum is 0.038, which rules out negative effects. Table 4 also presents 95 percent confidence intervals, which are based on the results by Imbens and Manski (2004) – since these bounds do not involve min or max operators – and asymptotically cover the true parameter value with 0.95 probability. As above, while the confidence intervals do not rule out negative effects under the monotonicity assumption, they imply a positive effect once the mean dominance assumption is added. This illustrates the identifying power of this additional assumption.
The second column of Table 4 presents results for Non-Hispanics, for whom the monotonicity assumption of $S$ in $Z$ is more plausible. In general, the lower and upper bounds under the two sets of assumptions are larger for Non-Hispanics than for the entire population. Under the monotonicity and mean dominance assumptions, the estimated lower and upper bounds on the ITT effect of JC on $\ln(\text{wage})$ for the $EE$ stratum are 0.050 and 0.119, respectively, and the 95 percent confidence interval is $[0.029, 0.144]$.

**Bounds on $\Delta$.** Table 5 shows the estimation results for the average effect of JC enrollment on $\ln(\text{wage})$ for always-employed compliers ($\Delta$ in (1)). This table shows the estimated stratum proportions, relevant quantities used in estimating the bounds, half-median unbiased estimates of our bounds, and CLR confidence intervals for the true parameter values under Assumptions 1 through 5 (Proposition 1) and under Assumptions 1 through 6 (Proposition 2). The first column presents the results for the entire sample, and the second shows the results for Non-Hispanics. For both samples, the largest stratum is the $cEE$ stratum, with an estimated proportion of almost 40 percent, while the stratum of always-employed always-takers ($aEE$) is the smallest one. All the estimated stratum proportions in Table 5 are statistically different from zero.

For the entire sample, the estimated lower and upper bounds on $\Delta$ under Proposition 1 are $-0.022$ and 0.130, respectively, while the corresponding numbers for Non-Hispanics are $-0.014$ and 0.161. Given the weak effects of JC on labor market outcomes for Hispanics, it is not surprising that the bounds for Non-Hispanics cover a larger positive region than those for the entire population. For both samples, the estimated lower and upper bounds are larger than the corresponding bounds for the ITT effect in Table 4, especially the upper bound (e.g., for Non-Hispanics, 0.119 versus 0.161, or a 35.3% increase). From Tables 4 and 5, the positive region covered by the bounds on the effect of JC enrollment on wages for the $cEE$ stratum is larger than the positive region covered by the bounds on the ITT effect of JC on wages for the $EE$ stratum. This suggests that the effects of JC on wages obtained by Lee (2009) were conservative, since the effect was weakened by noncompliance to the assigned treatment.

As in Lee (2009), we are unable to rule out a zero effect of JC on wages using only the monotonicity assumption of the effect of JC on employment. However, as before, our lower bounds are implicitly constructed under the implausible worst-case scenario of a perfect negative
Table 5: Bounds for the Effect of JC on ln(wage) for the cEE Stratum

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>Non-Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{a_{EE}}$</td>
<td>0.016** (.002)</td>
<td>0.018** (.002)</td>
</tr>
<tr>
<td>$\pi_{n_{EE}}$</td>
<td>0.158** (.005)</td>
<td>0.160** (.005)</td>
</tr>
<tr>
<td>$\pi_{c_{EE}}$</td>
<td>0.391** (.010)</td>
<td>0.381** (.011)</td>
</tr>
<tr>
<td>$\pi_{c_{NE}}$</td>
<td>0.041** (.011)</td>
<td>0.050** (.012)</td>
</tr>
<tr>
<td>$\pi_{a_{NN}}$</td>
<td>0.028** (.003)</td>
<td>0.030** (.003)</td>
</tr>
<tr>
<td>$\pi_{n_{NN}}$</td>
<td>0.104** (.004)</td>
<td>0.104** (.005)</td>
</tr>
<tr>
<td>$\pi_{c_{NN}}$</td>
<td>0.261** (.007)</td>
<td>0.258** (.008)</td>
</tr>
<tr>
<td>$\alpha_{c_{EE}}$</td>
<td>0.872** (.023)</td>
<td>0.849** (.025)</td>
</tr>
<tr>
<td>$E[Y^*(1)</td>
<td>a_{EE}]$</td>
<td>2.033** (.059)</td>
</tr>
<tr>
<td>$E[Y^*(0)</td>
<td>n_{EE}]$</td>
<td>2.033** (.016)</td>
</tr>
<tr>
<td>$E[Y^*(0)</td>
<td>c_{EE}]$</td>
<td>1.972** (.015)</td>
</tr>
<tr>
<td>$\nu[\gamma_{1_{c_{EE}}}^{11} \leq Y \leq \gamma_{1_{c_{NE}}}^{11}]$</td>
<td>2.429** (.066)</td>
<td>2.376** (.057)</td>
</tr>
<tr>
<td>$\nu[\gamma_{1_{c_{EE}}}^{11} \leq Y \leq \gamma_{1_{c_{EE}}}^{11}]$</td>
<td>1.676** (.034)</td>
<td>1.703** (.035)</td>
</tr>
</tbody>
</table>

**Notes:** 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 design weights. 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 $\alpha_{c_{EE}}$ and $\alpha_{c_{NE}}$ are provided in Proposition 1.
correlation between employment and wages, which is contradicted by standard economic models. The mean dominance assumption we use rules out this implausible extreme case and helps to increase the lower bound and to narrow the bounds. The last part of Table 5 shows results when the monotonicity and mean dominance assumptions are both used. In this case, the estimated lower bounds on the average effect of JC on \( \ln(\text{wage}) \) for the \( cEE \) stratum are 0.055 and 0.074 for the entire and Non-Hispanics samples, respectively, and the CLR-based 95 percent confidence intervals for the true value of \( \Delta \) exclude zero for both samples. Thus, under all six assumptions, our results imply positive average effects of JC on wages for the \( cEE \) stratum in both the entire and Non-Hispanics samples. These results illustrate the identifying power of Assumption 6, and also reinforce the notion that the ITT effects of JC on wages are likely to be lower than the effect of JC enrollment on wages. Finally, note that, as discussed in Remark 2, in the absence of Assumption 5 the lower bound for \( \Delta \) in Proposition 2 can be interpreted as the lower bound for \( \beta = E[Y^*(1) - Y^*(0)|cEE, cNE] \) under Assumptions 1 through 4, and 5'.

**Missing Values due to Non-Response.** To analyze the sensitivity of our main results to missing values of relevant variables, and to compare our results to those in Frumento et al. (2012), we estimate bounds on \( \Delta \) accounting for this issue. We do this by employing another weight constructed by the NJCS using non-public data that accounts for sample design, 48-month interview design, and 48-month interview non-response. The key assumption is that the probability that the information is missing for a given individual is random conditional on the set of variables used to construct the weight. Frumento et al. (2012) employ the same assumption but only use variables available in the public version of the NJCS data. To construct the data used in this exercise, we include all the individuals who responded to the 48-month interview and we drop those with missing values for weekly working hours, weekly earnings, or enrollment information. The estimation results for the bounds on \( \Delta \) are presented in Table 6 (the corresponding results for the other parameters shown in Table 5 are provided in the Internet Appendix). The estimated lower and upper bounds for the effect of JC on \( \ln(\text{wages}) \) for the \( cEE \) stratum are slightly below those in Table 5, which ignore the non-response issue. Focusing on Non-Hispanics, the positive region covered by the estimated bounds from Proposition 2 is slightly less than the region covered by the corresponding bounds in Table 5. In general, we
Table 6: Bounds for the Effect of JC for $cEE$ Adjusting for Non-Response

<table>
<thead>
<tr>
<th></th>
<th>Entire Sample</th>
<th>Non-Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{cEE}$</td>
<td>.391** (.009)</td>
<td>.381** (.010)</td>
</tr>
<tr>
<td>$\alpha_{cEE}$</td>
<td>.877** (.022)</td>
<td>.851** (.024)</td>
</tr>
</tbody>
</table>

Bounds under Monotonicity (Proposition 1):

$[LY_{1,cEE}, UY_{1,cEE}]$  
CLR 95% confidence interval  
CLR 95% confidence interval

$[L_{cEE}, U_{cEE}]$  
CLR 95% confidence interval  
CLR 95% confidence interval

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 $\alpha_{cEE}$ and $\alpha_{NE}$ are provided in Proposition 1.

conclude that, although adjusting for non-response slightly weakens our previous findings, the results still strongly suggest a positive average effect of JC enrollment on wages four years after random assignment for the always-employed compliers.

Finally, we compare these results to the ones in Frumento et al. (2012), who point identify the average effect of JC enrollment on wages, adjusting for sample selection, noncompliance, and missing outcomes, by imposing a different set of assumptions from the two sets we consider. Employing a different sample from ours, their results imply an effect on ln(wage) of about $0.038$ for the $cEE$ stratum. This point estimate is consistent with the estimated bounds for this effect presented in Table 6 for the entire population under our Assumptions 1 through 5 and, although it is below the estimated lower bound under Assumptions 1 through 6 ($0.044$), it is inside the corresponding 95 percent confidence interval. Thus, their point estimate of the effect of JC on wages for the $cEE$ stratum is consistent with our bounds adjusting for non-response.

3.4 Conclusions from Empirical Analysis

We draw the following conclusions from our empirical analysis. First, our results strongly suggest a positive average effect of participating in JC on wages four years after random assignment.
for the always-employed compliers. For Non-Hispanics, for whom the monotonicity assumption on the effect of JC on employment is more likely to hold, our estimated bounds under Assumptions 1 through 6 in Table 5 imply that participating in JC increases the average wage (as opposed to ln(wage)) of the always-employed compliers by between 7.7 and 17.5 percent. This evidence suggests that JC has an effect on participants’ earnings not only by increasing their probability of being employed but also by increasing their wages, which is most likely a consequence of their human capital accumulation during enrollment in JC.

Second, our analysis suggests that the results from the studies of the ITT effects of JC on wages in Lee (2009) and Blanco et al. (2013) are conservative because noncompliance is likely to dilute the effect of JC enrollment on wages. In particular, we find that for the two samples we consider, and regardless of whether or not we employ Assumption 6, the positive region covered by the bounds on the effect of JC enrollment on wages for the cEE stratum is larger than the positive region covered by the bounds on the ITT effect of JC on wages for the EE stratum. This is consistent with the results presented in Section 3.1, which show that the LATE estimates of the effects of JC on other labor market outcomes not suffering from sample selection are larger than the corresponding ITT estimates. This conclusion is also consistent with that by Frumento et al. (2012), who find that their point estimate of the effect of JC enrollment on wages at week 208 for the cEE stratum is larger than the point estimate of the corresponding ITT effect for the EE stratum in Zhang et al. (2009).

4 Simulation

This section presents a simulation study aimed at shedding light on the performance of our bounds when some of our assumptions are violated, and when all our assumptions hold. In particular, we use two simulation designs to analyze violations (or near violations) of: (1) Assumption 5 (no cEN stratum); (2) Assumption 6 (mean dominance); (3) both Assumptions 5 and 6; (4) Assumption 3 \( E[D(1) - D(0)] \neq 0 \); and, (5) Assumption 4 (no defiers).

In order for our simulation design to be close to situations found in empirical research, our designs mimic the characteristics of our entire sample (described in Section 3.1). Both of our designs share the following characteristics. Each simulated sample contains 9090 observations,
with each observation receiving a randomly assigned value of $Z = 1$ with probability 5491/9090. The individual’s membership to principal strata is drawn from a uniform distribution, where the fraction of each stratum in the population is explained below. The observed treatment and employment statuses, $D$ and $S$, are jointly determined by the membership to principal strata and $Z$. The individuals’ wages in each of the strata follow a lognormal distribution with variance equal to $\sigma^2$, and means equal to $E[Y^{*(1)}|cEE] = 2.04$, $E[Y^{*(0)}|cEE] = 1.97$, $E[Y^{*(1)}|aEE] = 2.035$, $E[Y^{*(0)}|nEE] = 2.035$, and $E[Y^{*(0)}|cEN] = 2$. Thus, the true effect $\Delta$ is equal to .07. Those parameter values are chosen to be close to their corresponding estimated values, and to our estimates of $\overline{Y}^{zds}$, based on our original sample (e.g., see Table 5).

The first design is used to analyze violations to Assumptions 5 and 6. It sets the true values of the stratum proportions to be close to their corresponding estimated values (and to our estimates of $p_{ds|z}$) based on our original sample as: $\pi_{aNN} = .030$, $\pi_{aEE} = .015$, $\pi_{nNN} = .105$, $\pi_{nEE} = .155$, $\pi_{cEE} = .395 - \pi_{cEN}$, $\pi_{cNN} = .26 - \pi_{cEN}$, and $\pi_{cNE} = .04 + \pi_{cEN}$. The total proportion of compliers ($\pi_c$) equals $E[D(1) - D(0)] = 0.695$, as in our original sample, and $\pi_{cEN} \neq 0$ without Assumption 5. Finally, the individuals’ wages in the $cEN$ stratum follow a lognormal distribution with mean $E[Y^{*(0)}|cEN] = 2$ and variance $.2$.

Simulation 1 considers violations to Assumption 5, while mean dominance holds with $E[Y^{*(1)}|cEE] - E[Y^{*(1)}|cNE] = .16$ (chosen to be consistent with our original estimate of $\overline{Y}^{111}$). Figure 1a shows the average of the estimated lower and upper bounds over 1000 replications for values of $\pi_{cEN}$ between 0 and .26 (outside this range some stratum proportions are negative). To give a sense about the variability of the estimated bounds, Figure 1a also reports the 95th percentile of the estimates of $U_{cEE}$ from the 1000 replications, and the corresponding 5th percentiles of the estimates of $L_{cEE}$ and $\overline{L}_{cEE}$. For example, 95 percent of the estimates of $L_{cEE}$ lie above the plotted squares, while 95 percent of the estimates of $U_{cEE}$ lie below the plotted circles. For clarity, we omit the 5th and 95th percentiles for the upper and lower bounds, respectively (they are provided in the Internet Appendix).

Figure 1a reports that the means of the estimated bounds decrease as $\pi_{cEN}$ increases. This happens because the estimates of $E[Y^{*(0)}|cEE]$ increase while the estimated bounds on $E[Y^{*(1)}|cEE]$ decrease. As $\pi_{cEN}$ increases, the fraction of $cEE$ individuals in the cell
\( \{Z = 0, D = 0, S = 1\} \) falls while the fraction of \( cEN \) individuals increases. Since in our design the mean of \( Y^*(0) \) for the \( cEN \) stratum is larger than that of the \( cEE \) stratum, \( \bar{Y}^{001} \) and thus our estimate of \( E[Y^*(0)|cEE] \) increase (see equation 4). Something similar occurs in the cell \( \{Z = 1, D = 1, S = 1\} \), where the fraction of \( cEE \) individuals decreases while the fraction of \( cNE \) individuals increases (since \( \pi_{NE} = .04 + \pi_{cEN} \)), which combined with the mean of \( Y^*(1) \) being larger for the former than the latter individuals results in smaller bounds for \( E[Y^*(1)|cEE] \). The true effect is above the mean estimates of the lower bounds over the entire range of \( \pi_{cEN} \), and is above the mean estimates of \( U_{cEE} \) when \( \pi_{cEN} > .16 \).

Figure 1b shows the percentage of times out of the 1000 repetitions that the true effect is within the estimated bounds. For Proposition 1, this number gradually declines as \( \pi_{cEN} \) increases. For Proposition 2, it starts lower than for Proposition 1 because the true value of \( \bar{L}_{cEE} \) when \( \pi_{cEN} = 0 \) is close to \( \Delta \). The estimated bounds cover \( \Delta \) in both propositions at least 80 percent of the time when \( \pi_{cEN} < .1 \). When \( \pi_{cEN} = 0 \) and all our assumptions hold, the means of the simulated bounds are \([-.027, .136]\) in Proposition 1 and \([.054, .136]\) in Proposition 2, which are similar to the empirical results for the entire sample in Table 5.

Simulation 2 considers violations of Assumption 6 when \( \pi_{cEN} = 0 \). Figure 2a presents results for different values of \( E[Y^*(1)|cEE] - E[Y^*(1)|cNE] \), where the mean dominance is violated for
negative values of this difference. As this difference increases, the means of the estimated bounds fall because larger values of the $cNE$ members’ outcomes in the cell $\{Z = 1, D = 1, S = 1\}$ are being replaced with smaller values, which results in smaller estimated bounds on $E[Y^*(1) | cEE]$ (here the stratum proportions and $E[Y^*(1) | cEE]$ are held fixed). While the true effect is within the mean estimated bounds in Proposition 1 over the entire range considered, it is below the mean of the estimates of $\bar{T}_{cEE}$ when Assumption 6 is violated. The reason is that, by construction, if this assumption is violated then $E[Y^*(1) | cNE]$ provides an upper (rather than a lower) bound for $E[Y^*(1) | cEE]$, and thus $\bar{T}_{cEE}$ provides an upper bound for $\Delta$. Figure 2b reports that $\Delta$ is within the estimated bounds in Proposition 1 in almost every replication. In contrast, the percentage of times $\Delta$ is within the estimated bounds in Proposition 2 increases gradually as the values of $E[Y^*(1) | cEE] - E[Y^*(1) | cNE]$ increase. At zero, the percentage is about 50%. However, note that the 5th percentile of the estimates of $\bar{T}_{cEE}$ is below the true value of $\Delta$ for an important range of negative values of that difference.

Simulation 3 considers the case when both Assumptions 5 and 6 fail. Figure 3 shows plots similar to those in Figure 2 when $\pi_{cEN}$ equals .05, .10 and .15. As $\pi_{cEN}$ increases in Figures 3a, 3b, and 3c, $\Delta$ is within the mean estimated bounds in Proposition 1 over a shrinking range of $E[Y^*(1) | cEE] - E[Y^*(1) | cNE]$. The intersection of the mean of the estimates of $\bar{T}_{cEE}$ with $\Delta$ moves slightly leftward as $\pi_{cEN}$ increases. In contrast to Figure 2b, Figures 3d, 3e, and
3f display a bell shape of the percentage of times $\Delta$ is within the estimated bounds, with the bell covering a shrinking area as $\pi_c$ increases. This implies that when Assumption 5 is violated, larger values of $E[Y^*(1)|cEE] - E[Y^*(1)|cNE]$ tend to have a lower probability that the estimated bounds cover the true effect. The percentages for Proposition 1 in those figures are above those for Proposition 2, with the bell shape being centered at zero. Instead, the center for Proposition 2 is in the positive region and moves slightly left as $\pi_c$ increases.

Next we examine the performance of our bounds when the commonly used IV Assumptions 3 and 4 are violated or close to be violated. To be able to vary the proportion of defiers ($\pi_d$) and the strength of the IV (i.e., the value of $E[D(1) - D(0)]$), we use a second design. Let $\pi_n$ and $\pi_a$ denote the proportions of never and always takers, respectively. Based on our original sample, the ratios between different strata are set as follows: $\pi_n/\pi_a = 5.7$, $\pi_{aEE}/\pi_a = .3$, $\pi_{nEE}/\pi_n = .6$, $\pi_{cEE}/\pi_c = .559$, $\pi_{cNE}/\pi_c = .059$, $\pi_{dEE}/\pi_d = .5$, and $\pi_{dEN}/\pi_d = .1$. The specific stratum proportions are obtained by setting the values of $E[D(1) - D(0)]$ and $\pi_d$ (since $\pi_c = E[D(1) - D(0)] + \pi_d$, where Assumption 5 implies $\pi_{dNE} = 0$. The individuals’s wages in the strata follow a lognormal distribution with variance $.2$ and means: $E[Y^*(1)|dEE] = 2.02,$
$E[Y^*(0)|dEE] = 1.99$, $E[Y^*(1)|dEN] = 1.85$, and $E[Y^*(1)|cNE] = 1.88$ (the means of the other strata are the same as before).

Note that, since we use IV methods to address noncompliance, our bounds are likely to be affected by a weak IV, i.e., one for which $E[D(1) - D(0)]$ is close to 0. The reason is that the term $\pi_{cEE} = p_{01|0} - p_{01|1}$, which is no greater than $E[D(1) - D(0)]$, appears in the denominator in several of the expressions in Propositions 1 and 2 (see also, e.g., equation 4), and thus values of $p_{01|0} - p_{01|1}$ close to 0 can lead to large variance and instability in our estimated bounds. The trimming proportions $\alpha_{cEE}$ and $\alpha_{cNE}$ are also negatively affected by small values of $\pi_{cEE}$ and $\pi_{cNE}$, leading to wider trimming bounds. Moreover, weak IVs are known to exacerbate the bias of IV estimators coming from the presence of defiers (AIR). In our setting, we can write the right hand side of equation (4) as $E[Y^*(0)|cEE]$ plus a bias term equal to (details provided in the Internet Appendix): $\pi_{dEE}(E[Y^*(0)|cEE] - E[Y^*(0)|dEE])/(\pi_{cEE} - \pi_{dEE})$. A weak IV implies a small value of the denominator and thus a greater bias coming from the presence of defiers. As a result, our simulation when Assumption 4 fails considers different IV strengths.

Simulation 4 analyzes cases when Assumption 4 holds but Assumption 3 is violated or close to be violated. Figure 4 shows results when we vary the proportion of compliers ($\pi_c$), which equals $E[D(1) - D(0)]$ when $\pi_d = 0$, from 0 to .7. This range covers the estimated value of $\pi_c$ in our original sample (.695). Because of the large variation in our estimated bounds when $\pi_c$ is very close to zero, we present the results in two sets of graphs. The first set varies $\pi_c$ from 0 to .05, and the second varies $\pi_c$ from .05 to .7 (note the change of scale in the y-axis of those figures). For reference, we refer to these two cases as “weak IV” and “stronger IV”, respectively. Figure 4a shows that there is substantial variation and instability in our estimated bounds when $\pi_c$ is close to zero. Moreover, the closer $\pi_c$ is to zero, the more likely we are to have replications in which the estimated lower bound is above the estimated upper bound or to have some negative estimated stratum proportions (not shown in figures). The poor performance of our bounds in such cases is also reflected in Figure 4c with a very small probability that the estimated bounds cover the true effect. Importantly, Figures 4b and 4d show that, as $\pi_c$ moves away from zero, the performance of our bounds improve significantly. Note also from Figures 4c and 4d that, except for very low values of $\pi_c$, the percentage of times $\Delta$ is within the estimated bounds
bounds in Proposition 1 is greater than that for the estimated bounds in Proposition 2.

Simulation 5 analyzes the case when Assumption 4 is violated for different values of $E[D(1) - D(0)]$: .695, .4, and .02. For reference, we label these cases strong, moderate, and weak IV, respectively. Figure 5 shows results for those three cases as we vary $\pi_d$ (whose possible values depend on $E[D(1) - D(0)]$). Figure 5a reports that the true effect is within the mean estimated bounds in Proposition 2 for $\pi_d \leq .15$ with a strong IV, and for $\pi_d \leq .1$ with a moderate IV. For the strong and moderate IV cases, the true effect is within the mean estimated bounds in Proposition 1 for all the values of $\pi_d$ considered in each case. As before, Figure 5c shows that there is a lot of instability in our bounds when the IV is weak. The graphs at the bottom of Figure 5 show that the percentage of times the estimated bounds cover the true effect decreases.
as $\pi_d$ increases. They also illustrate the importance that the strength of the IV has on the performance of our bounds, as the percentages decrease dramatically as the IV becomes weaker.

In general, we conclude from the simulation exercises above that our bounds seem to be relatively robust to small violations of Assumptions 5 and 6. Our results also imply that the strength of the IV affects the sensitivity of our bounds to Assumption 4: with a strong IV (which is the case in our JC application) our bounds are relatively robust to the presence of defiers. The simulations also illustrate an important trade-off between the bounds in Propositions 1 and 2: while the estimated bounds in Proposition 2 are narrower than those in Proposition 1, the probability that the estimated bounds in Proposition 2 cover the true effect is usually lower, especially when some of our assumptions are violated. This is an important consideration when employing our bounds in practice. Finally, our results show that our bounds perform poorly when the IV is weak, thus warning researchers against their use in such situations. This is not surprising given the poor performance of standard IV estimators with weak IVs (e.g., Staiger and Stock, 1997).
5 Conclusion

We derive nonparametric bounds for average treatment effects in the presence of both sample selection and noncompliance under relatively weak assumptions, and employ these bounds to empirically assess the wage effects of participating in Job Corps (JC). The results from a simulation study suggest that our bounds are relatively robust to small violations of our assumptions, but warn against their use when the instrument employed to address noncompliance is weak.

The first contribution of the paper is to extend the bounds derived in Zhang et al. (2008) and Lee (2009), which address only sample selection, to settings where noncompliance is also present. More generally, our bounds can be used in settings where two identification issues are present and there is an IV to address one of them. For example, our methods could be applied to analyze time-to-event outcomes (which are usually right censored) when another identification issue is present. In the context of JC, an important question is whether enrollment in JC has an effect on its participants’ unemployment spells, which are right censored because not all individuals find employment before the end of the study. In this case, our methods would provide information about the effect of JC on unemployment spells for those compliers who would always find a job before the end of the study whether or not they enrolled in JC. Moreover, the bounds developed in Zhang et al. (2008) and Lee (2009) are becoming increasingly used in the partial identification literature to address other problems. For instance, Huber and Mellace (2011) employ those bounds to test the implications of the exclusion restriction assumption in just-identified models, while Flores and Flores-Lagunes (2010) use those bounds to construct bounds on direct and indirect effects. The bounds derived in this paper can be employed to extend the results in those papers to address more than one complication. In addition, the methods herein can be combined with those in Huber and Mellace (2013) to construct bounds on the average effect for subpopulations other than the $cEE$ stratum (for example, for the treated and selected individuals, who have $D = 1$ and $S = 1$).

The second contribution of the paper is to study the wage effects of JC. Our results strongly suggest a positive average effect of participating in JC on wages four years after randomization for the always-employed compliers. The results also point to larger positive wage effects of
JC than those found without adjusting for noncompliance in Lee (2009) and Blanco et al. (2013). Our findings suggest that JC has positive effects not only on the employability of its participants but also on their wages, implying that JC is likely to have positive effects in their human capital. Therefore, it is very important to consider the potential benefits of JC and other training programs on wages when evaluating their effects.

6 Acknowledgments

Detailed comments from the Editor, Associated Editor, and three anonymous referees greatly improved the paper and are gratefully acknowledged. We are also grateful for comments from Marianne Bitler, Eduardo Fajnzylber, Alfonso Flores-Lagunes, Laura Giuliano, Guido Imbens, Fabrizia Mealli, Oscar Mitnik, Christopher Parmeter, Jeffrey Smith, Zhong Zhao, James Ziliak, and conference/seminar participants at University of Miami, Renmin University of China, the Theory and Practice of Program Evaluation Workshop at CEPS/INSTEAD, the 2012 Impact Evaluation Network Meeting at Harvard University, the 2012 Midwest Econometrics Group Meeting at University of Kentucky, the 2014 Southern California Conference in Applied Microeconomics at Claremont McKenna College, and the 2014 International Association for Applied Econometrics Conference at Queen Mary University of London. Flores acknowledges summer research support from the Orfalea College of Business at California Polytechnic State University. All errors are our own.

References


