DISCUSSION PAPER SERIES

IZA DP No. 9511

Going Beyond *LATE*: Bounding Average Treatment Effects of Job Corps Training

Xuan Chen
Carlos A. Flores
Alfonso Flores-Lagunes

November 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Going Beyond *LATE*: Bounding Average Treatment Effects of Job Corps Training

Xuan Chen
Renmin University of China

Carlos A. Flores
California Polytechnic State University

Alfonso Flores-Lagunes
Syracuse University and IZA

Discussion Paper No. 9511
November 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 9511
November 2015

ABSTRACT

Going Beyond *LATE*: Bounding Average Treatment Effects of Job Corps Training*

We derive nonparametric sharp bounds on average treatment effects with an instrumental variable (IV) and use them to evaluate the effectiveness of the Job Corps (JC) training program for disadvantaged youth. We concentrate on the population average treatment effect (*ATE*) and the average treatment effect on the treated (*ATT*), which are parameters not point identified with an IV under heterogeneous treatment effects. The main assumptions employed to bound the *ATE* and *ATT* are monotonicity in the treatment of the average outcomes of specified subpopulations, and mean dominance assumptions across the potential outcomes of these subpopulations. Importantly, the direction of the mean dominance assumptions can be informed from data, and some of our bounds do not require an outcome with bounded support. We employ these bounds to assess the effectiveness of the JC program using data from a randomized social experiment with non-compliance (a common feature of social experiments). Our empirical results indicate that the effect of JC on eligible applicants (the target population) four years after randomization is to increase weekly earnings and employment by at least $24.61 and 4.3 percentage points, respectively, and to decrease yearly dependence on public welfare benefits by at least $84.29. Furthermore, the effect of JC on participants (the treated population) is to increase weekly earnings by between $28.67 and $43.47, increase employment by between 4.9 and 9.3 percentage points, and decrease public benefits received by between $108.72 and $140.29. Our results also point to positive average effects of JC on the labor market outcomes of those individuals who decide not to enroll in JC regardless of their treatment assignment (the so-called never takers), suggesting that these individuals would indeed benefit from participating in JC.

JEL Classification: J30, C13, C21

Keywords: training programs, program evaluation, average treatment effects, bounds, instrumental variables

Corresponding author:

Alfonso Flores-Lagunes
Department of Economics and Center for Policy Research
Syracuse University
426 Eggers Hall
Syracuse, NY 13244-1020
USA
E-mail: afloresl@maxwell.syr.edu

* We are grateful for comments from Joshua Angrist, Wallice Ao, Dan Black, Ying-Ying Lee, Ismael Mourifié, Jeff Smith, and seminar participants at University of Miami, California Polytechnic State University at San Luis Obispo, University of Central Florida, Queens College (CUNY), the 2012 New York Camp Econometrics, the 2012 Midwest Econometrics Group Meetings at University of Kentucky, the 2014 Annual Meetings of Society of Labor Economists, the 13th IZA/SOLE Transatlantic Meeting of Labor Economists, the 2014 California Econometrics Conference at Stanford University, and the 2014 Annual Meetings of the Southern Economic Association. Flores acknowledges funding from the Research, Scholarship, and Creative Activities Grant program at California Polytechnic State University. Previous versions of this paper circulated under the title "Bounds on Population Average Treatment Effects with an Instrumental Variable." All the usual disclaimers apply.

1 Introduction

Government-sponsored training programs are essential tools to help improve the labor market prospects of economically disadvantaged citizens and reduce their dependence on safety net programs. As such, the evaluation of the effectiveness of training programs is a critical issue that has generated a large empirical and methodological literature (e.g., Lalonde, 1986; Dehejia and Wahba, 1999; Heckman et al., 1999). In the United States, Job Corps (JC) is the main training program targeted to disadvantaged youth. It delivers a comprehensive bundle of benefits to approximately 61,000 participants a year at a cost of about \$1.5 billion (US Department of Labor, 2015). In order to evaluate the effectiveness of this large-scale training program, the United States Congress authorized the National Job Corps Study (NJCS), a randomized social experiment. The randomized nature of the NJCS was intended to provide uncontroversial findings given its reliance on weak assumptions relative to other evaluation methods (e.g., LaLonde, 1986; Heckman et al., 1999). Nevertheless, the NJCS was subject to non-compliance (e.g., only about 73 % of treatment-group individuals enrolled in JC). Under non-compliance, researchers typically focus on the "intention-to-treat" (ITT) effect that takes the randomization as the treatment of interest, or on the "local average treatment effect" ($LATE$) that corresponds to the effect of the training program for a particular subset of individuals. Both of the previous effects fall short of the average effect of the training program for the population or for those undergoing training—parameters of first order importance in the evaluation literature (e.g., Heckman et al., 1999). To the best of our knowledge, there are no estimates of the latter parameters using data from the NJCS. We fill this gap by undertaking inference on them.

Estimation of the $LATE$ in experiments where subjects do not comply with their randomized treatment assignment is accomplished by using the treatment assignment indicator as an instrumental variable for the actual treatment receipt indicator. Instrumental variable (IV) methods have been widely used in the literature of program evaluation due to its high internal validity. An influential framework for studying causality using IVs was developed by Imbens and Angrist (1994), and Angrist, Imbens and Rubin (1996) (hereafter IA and AIR, respectively). They show that, in the presence of heterogeneous effects, IV estimators point identify the local average treatment effect ($LATE$) for compliers, a subpopulation whose treatment status is affected by the instrument. Common criticisms of their framework are the focus on the effect for a subpopulation and the instrument-specific interpretation of the $LATE$ (e.g., Heckman, 1996; Robins and Greenland, 1996; Deaton, 2010; Heckman and Urzua, 2010). As a result, a growing literature pursues the external validity of IV methods. Point identification of population treatment effects usually requires an instrument to be strong enough to drive the probability of being treated from zero to one (e.g., Heckman, 2010), which is hard to satisfy in practice. Another strategy relies on stable IV estimates revealed empirically, which inspire the use of multiple instruments for the same causal relationship (e.g., Angrist and Fernandez-Val, 2010). Unfortunately, finding multiple IVs can be challenging in practice.

An alternative to point identification of treatment effects other than $LATE$ using IVs is partial identification. Manski (1990) pioneered partial identification of the population average treatment effect (ATE) under the mean independence assumption of the IV. Since then, there has been a growing literature on partial identification of the ATE with IV methods. One strand of this literature endeavors to improve Manski's (1990) bounds by imposing different monotonicity assumptions. Manski (1997) derived bounds under the monotone treatment response (MTR) assumption, which asserts monotonicity of the outcome in the treatment. Manski and Pepper (2000) introduced the monotone instrumental variable (MIV) assumption, which states that mean response varies weakly monotonically across subpopulations with different levels of the instrument (as opposed to be constant, like in the traditional mean independence of the IV assumption). Chiburis (2010a) added the mean independence of the IV assumption to both the MTR assumption and a special case of the MIV assumption to derive bounds on ATE that do not require specifying the direction of the monotonicity a priori. Another strand of the partial identification literature employs structural models on the treatment or the outcome to derive bounds. For instance, under the statistical independence of the IV assumption, Heckman and Vytlacil (2000) imposed a threshold crossing model with a separable error on the treatment. Focusing on a binary outcome, Shaikh and Vytlacil (2011) imposed threshold crossing models on both the treatment and the outcome; while Chiburis (2010b) considered a threshold crossing model on the outcome. Instead of assuming a threshold crossing model with separable errors, Chesher (2010) derived bounds by imposing a non-separable structural model on the outcome and assuming the structural function is weakly increasing in the non-separable error.

Given the alternative assumptions for partial identification of the ATE with IVs, a comparison of their identification power is important. First, the monotonicity assumption of the treatment in the IV (e.g., IA; AIR; Balke and Pearl, 1997; Huber and Mellace, 2010) and the structural model assumptions on the treatment (e.g., Heckman and Vytlacil, 2000) do not improve on the informational content (i.e., width) of Manski's bounds derived under the mean independence of the IV assumption.[1] Second, monotonicity assumptions of the outcome in the treatment (e.g., Manski, 1997; Manski and Pepper, 2000) and the structural model assumptions on the outcome (e.g., Bhattacharya, Shaikh and Vytlacil, 2008, hereafter BSV; Chiburis, 2010a, 2010b; Chesher 2010; Shaikh and Vytlacil, 2011) do improve on Manski's bounds. Third, partial identification with IV methods usually requires bounded support of the outcome, which is a reason why most papers focus on binary outcomes (e.g., Balke and Pearl, 1997; BSV; Hahn, 2010; Chiburis, 2010b; Shaikh and Vytlacil, 2011). It is worth noting that for the case of a binary outcome several of the assumptions (and bounds) are equivalent. For example, Machado et al. (2009) showed the equivalence between the MTR assumption and the threshold crossing model on the outcome, while BSV showed that, in the absence of covariates, the bounds for a binary outcome under the MTR and mean independence of the IV assumptions are equivalent to those derived using threshold crossing models on both the treatment and the outcome.

This paper contributes to two different literatures. First, it contributes to the partial identification

[1] Vytlacil (2002) shows that the assumptions of independence and monotonicity of the IV on the treatment in the $LATE$ approach are equivalent to those of structural threshold crossing models on the treatment.

literature by deriving nonparametric sharp bounds for the ATE and the average treatment effect on the treated (ATT) by extending the work of IA and AIR. The proposed bounds improve on Manski's (1990) bounds and, importantly, some of them do not require a bounded outcome. We consider the setting of a binary instrument and a binary treatment, which is common in the existing literature on partial identification of treatment effects with IV methods. We contribute to the methodological literature two different sets of assumptions. The first is monotonicity in the treatment of the average outcomes of principal strata, which are subpopulations defined by the joint potential values of the treatment status under each value of the instrument. Similar to BSV and Shaikh and Vytlacil (2011), we do not require prior knowledge about the direction of the monotonicity. However, in contrast to the existing literature (e.g., Manski and Pepper, 2000; BSV; Shaikh and Vytlacil, 2011), we impose monotonicity on the average outcomes of strata rather than on the outcome of each individual. This is important as it makes the assumption more plausible in practice by allowing some individuals to experience a treatment effect that has the opposite sign to the ATE or ATT. The second set of assumptions involves mean dominance assumptions across the potential outcomes of different strata, which have been shown to have significant identifying power in other settings (e.g., Zhang et al., 2008; Flores and Flores-Lagunes, 2010, 2013; Chen and Flores, 2014; Huber et al., 2015). We propose to inform the direction of these mean dominance assumptions by comparing average baseline characteristics across strata that are likely to be highly correlated with the outcome.

In concurrent work to ours, Huber and Mellace (2010) and Huber et al. (2015) also derive nonparametric sharp bounds on average treatment effects within the $LATE$ framework. While both sets of work employ principal strata and consider mean dominance assumptions across these subpopulations, there are important differences between them. We consider the assumption of monotonicity in the treatment of the average outcomes of principal strata, which contains identifying power (thus narrowing the bounds) and can be justified by economic theory in certain applications. Furthermore, we consider additional variants of the mean dominance assumption across strata. On the other hand, we impose on our bounds the assumption of monotonicity of the treatment in the instrument, while those papers also consider bounds that do not impose this assumption.[2]

The second literature this paper contributes to is to that evaluating the effectiveness of JC, the largest federally-funded job training program for disadvantaged youth in the United States. Due to non-compliance, most studies evaluating JC using data from the NJCS concentrate on ITT effects or on the $LATE$ for individuals who comply with their random assignment (e.g., Schochet et al., 2001; Schochet et al., 2008; Flores-Lagunes et al., 2010). To the best of our knowledge, this is the first study that assesses the effectiveness of JC for eligible applicants (the target population) and program participants (the treated population) on three important outcomes: weekly earnings, employment, and the yearly amount of public benefits received. To this end, we employ the bounds on the ATE and the ATT derived herein.

[2]In general, estimated bounds without the assumption of monotonicity of the treatment in the instrument are wide in practice (Zhang et al., 2008; Blanco et al., 2013; Huber et al., 2015).

Using randomization into the program as an instrument for JC participation, the narrowest estimated bounds on the ATE four years after randomization derived under our assumptions are [\$24.61, \$201.04] for weekly earnings, [.042, .163] for employment, and [$-$\$142.76, $-$\$84.29] for public benefits. These results imply that the average effect of JC participation for eligible applicants is an increase of at least 11.6% and 7.2% on weekly earnings and employment, respectively, and a decrease of at least 9.9% in yearly dependence on public benefits. As compared to other bounds in the literature, those estimated bounds are significantly narrower than the estimated IV bounds proposed by Manski (1990), Heckman and Vytlacil (2000), and Kitagawa (2009) when applied to our setting, and the ones by Huber and Mellace (2010). Those estimated bounds are also narrower than those under the combination of the mean independence of the IV and MTR assumptions in Manski and Pepper (2000)—especially for public benefits—as well as those under the previous two assumptions plus a special case of the MIV assumption in Chiburis (2010a). Our estimated bounds on employment are also narrower than the ones proposed by Balke and Pearl (1997), BSV, Chesher (2010), Chiburis (2010b), and Shaikh and Vytlacil (2011) for the case of a binary outcome. The estimated bounds on the average effects of JC on participants (ATT) are substantially narrower than those on the ATE, providing a very tight interval where the true value of this effect lies. The narrowest estimated bounds for the ATT under our assumptions are [\$28.67, \$43.47] (about [13.5%, 20.4%]) for weekly earnings, [.049, .093] (about [8.4%, 16%]) for employment, and [$-$\$140.29, $-$\$108.72] (about [$-$16.5%, $-$12.8%]) for public benefits. In sum, our results indicate that JC has significant effects on the three outcomes analyzed, both for the population of eligible applicants (ATE) and for program participants (ATT). Importantly, estimated bounds that do not assume the sign of the average effect of JC on the outcomes for specific subpopulations are able to statistically rule out zero or negative ATEs and ATTs for weekly earnings and employment, as their 95 percent confidence intervals exclude zero.

Finally, as a by-product of our analysis, we also estimate bounds on the effects of JC participation for different strata. Our results are particularly informative for the stratum comprised of individuals who choose to never enroll in JC regardless of their treatment assignment (the so-called never takers). This is a key stratum from a policy perspective because these individuals are part of the target population of JC but decide against enrolling in it, probably believing they would not benefit from it. In our application, slightly more than one out of every four individuals belongs to this stratum. Thus, it is critical to determine whether those individuals would benefit, on average, from participating in JC. Our estimated bounds provide statistically significant evidence that, indeed, the average labor market outcomes of these individuals would be improved by participating in JC. More specifically, without imposing assumptions on the sign of the effects for this stratum, we find that their average weekly earnings and employment four years after randomization would be improved by at least \$13.03 (5.8%) and 2.5 percentage points (4.2%), respectively.

The rest of the paper is organized as follows. Section 2 presents the setup and the partial identification results on the ATE and ATT, with proofs relegated to the Appendix. Section 3 employs those bounds to analyze the effectiveness of the JC program, while Section 4 concludes.

2 Bounds on Average Treatment Effects

2.1 Setup and Benchmark Bounds

Consider a random sample of size n from a population. Let $D_i \in \{0,1\}$ indicate whether unit i is treated ($D_i = 1$) or not ($D_i = 0$), and let $Z_i \in \{0,1\}$ be an instrument for treatment. In our case, Z_i represents individual i's assignment to enroll ($Z_i = 1$) or not ($Z_i = 0$) in JC, while D_i represents her actual enrollment. Let $D_i(1)$ and $D_i(0)$ denote the treatment individual i would receive if $Z_i = 1$ or $Z_i = 0$, respectively. Let Y be the outcome (e.g., weekly earnings), and denote by $Y_i(1)$ and $Y_i(0)$ individual i's potential outcomes under treatment $D = d$, i.e., the outcomes individual i would experience if she received the treatment or not, respectively. Finally, let $Y_i(z, d)$ be the potential outcome as a function of the instrument and the treatment. Our parameters of interest are the population average treatment effect, $ATE = E[Y_i(1) - Y_i(0)]$, and the average treatment effect on the treated, $ATT = E[Y_i(1) - Y_i(0)|D_i = 1]$. For each unit, we observe $\{Z_i, D_i(Z_i), Y_i(Z_i, D_i(Z_i))\}$. This setting has received considerable attention in the literature (e.g., AIR, BSV). In what follows, we omit the subscript i unless necessary for clarity.

AIR partition the population into four strata based on the values of $\{D_i(0), D_i(1)\}$: $\{1,1\}$, $\{0,0\}$, $\{0,1\}$ and $\{1,0\}$. AIR (and the subsequent literature) refer to these strata as always takers (at), never takers (nt), compliers (c), and defiers (d), respectively. AIR impose the following assumptions, which we adopt hereafter:

Assumption 1 (Randomized Instrument). $\{Y(0,0), Y(0,1), Y(1,0), Y(1,1),\ D(0), D(1)\}$ is independent of Z.

Assumption 2 (Exclusion Restriction). $Y_i(0,d) = Y_i(1,d) = Y_i(d), d \in \{0,1\}$ for all i.

Assumption 3 (Nonzero First Stage). $E[D(1) - D(0)] \neq 0$.

Assumption 4 (Individual-Level Monotonicity of D in Z). Either $D_i(1) \geq D_i(0)$ for all i, or $D_i(1) \leq D_i(0)$ for all i.

Assumptions 1 through 3 are standard assumptions in the IV literature (e.g., IA, AIR). Assumption 4 rules out the existence of defiers (compliers) when the monotonicity is non-decreasing (non-increasing). The direction of the monotonicity can be inferred from the data given the independence of Z. Following BSV, we order Z so that $E[D|Z = 1] \geq E[D|Z = 0]$ to simplify notation in the rest of this section.[3]

Let $LATE_k = E[Y(1) - Y(0)|k]$ and π_k denote, respectively, the local (i.e., stratum-specific) average treatment effect and the stratum proportion in the population for stratum k, with $k = at$, nt, c. Let $\overline{Y}^{zd} = E[Y|Z = z, D = d]$ and $p_{d|z} = \Pr(D = d|Z = z)$. Under Assumptions 1 to

[3] It is possible to derive bounds on ATE and ATT without Assumption 4 (e.g., Huber et al., 2015) but in practice the resulting bounds are typically wide and unable to rule out zero (e.g., Zhang et al., 2008; Blanco et al., 2013; Huber et al., 2015). Therefore, we do not consider bounds without Assumption 4 in this paper.

4, the following quantities are point identified (IA; AIR): $\pi_{at} = p_{1|0}$, $\pi_{nt} = p_{0|1}$, $\pi_c = p_{1|1} - p_{1|0}$, $E[Y(1)|at] = \overline{Y}^{01}$, $E[Y(0)|nt] = \overline{Y}^{10}$ and $LATE_c = (E[Y|Z=1] - E[Y|Z=0])/(p_{1|1} - p_{1|0})$. Thus, in this setting the conventional IV estimand point identifies $LATE_c$, the local average treatment effect for compliers—units whose treatment status is affected by the instrument.[4]

We start by partially identifying the ATE. To this end, we write it as a function of the $LATE$s for always takers, never takers, and compliers:

$$
\begin{aligned}
ATE &= \pi_{at} LATE_{at} + \pi_{nt} LATE_{nt} + \pi_c LATE_c & (1) \\
&= p_{1|1}\overline{Y}^{11} - p_{0|0}\overline{Y}^{00} + p_{0|1}E[Y(1)|nt] - p_{1|0}E[Y(0)|at]; & (2)
\end{aligned}
$$

where $E[Y|Z=z] = E[E[Y|Z=z, D=d]|Z=z]$ is used in the second line.

By equation (2), since $Y(1)$ for never takers and $Y(0)$ for always takers are never observed in the data, additional assumptions are needed to bound the ATE. The most basic assumption considered in the previous literature (e.g., Manski, 1990) is the bounded support of the outcome.

Assumption 5 (Bounded Outcome). $Y(0), Y(1) \in [y^l, y^u]$.

This assumption states that the potential outcomes under the two treatment arms have bounded support. Replacing $E[Y(1)|nt]$ and $E[Y(0)|at]$ in equation (2) with either y^l or y^u, sharp bounds on the ATE under Assumptions 1 through 5 can be obtained.

Proposition 1 *Under Assumptions 1 through 5 the bounds $LB \leq ATE \leq UB$ are sharp, where*

$$
\begin{aligned}
LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^l p_{0|1} - y^u p_{1|0} \\
UB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}.
\end{aligned}
$$

The bounds in Proposition 1 are given here for reference since they represent a natural benchmark for the subsequent results. These bounds on the ATE coincide with the IV bounds in Manski (1990), Heckman and Vytlacil (2000), and Kitagawa (2009) when applied to the present setting; and with those in Huber et al. (2015). When the outcome is binary, these bounds also coincide with those in Balke and Pearl (1997).

2.2 Bounds on the ATE under Weak Monotonicity of Local Average Outcomes in the Treatment

The following is the first set of assumptions we consider to improve the identification power of the bounds in Proposition 1.

[4]Point identification of the rest of the quantities follows from Assumptions 1 and 4, as the latter implies that those observations with $\{Z=0, D=1\}$ are always takers, and those with $\{Z=1, D=0\}$ are never takers. For completeness, note that observations with $\{Z=0, D=0\}$ are either never takers or compliers, while those with with $\{Z=1, D=1\}$ are either always takers or compliers.

Assumption 6 (Weak Monotonicity in D of Average Outcomes of Strata). (i) Either $E[Y(1)|k] \geq E[Y(0)|k]$ for all $k = at, nt, c$; or $E[Y(1)|k] \leq E[Y(0)|k]$ for all $k = at, nt, c$. (ii) $E[Y(1)-Y(0)|c] \neq 0$.

Assumption 6(i) requires that the $LATE$s of the three existing strata are all either non-negative or non-positive. This assumption is similar to that in BSV, with the important distinction that we impose weak monotonicity on the $LATE$s rather than on the individual effects, which renders our assumption more plausible in practice by allowing some individuals to have a treatment effect of opposite sign to that of the ATE. Assumption 6(ii) is used to identify the direction of the monotonicity from the sign of the IV estimand ($LATE_c$) under the current assumptions. Note that, since we ordered Z so that $E[D|Z=1] \geq E[D|Z=0]$ (i.e., $p_{1|1} - p_{1|0} \geq 0$), the ITT effect $E[Y|Z=1] - E[Y|Z=0]$ and $LATE_c$ share the same sign. The following proposition presents sharp bounds on the ATE under the additional Assumption 6.

Proposition 2 *Under Assumptions 1 through 6 the bounds $LB \leq ATE \leq UB$ are sharp, where, if $E[Y|Z=1] - E[Y|Z=0] > 0$,*

$$\begin{aligned} LB &= E[Y|Z=1] - E[Y|Z=0] \\ UB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$,

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^l p_{0|1} - y^u p_{1|0} \\ UB &= E[Y|Z=1] - E[Y|Z=0]. \end{aligned}$$

Depending on the sign of $LATE_c$, either the lower or the upper bound in Proposition 2 improves upon the corresponding bound in Proposition 1. If $LATE_c > 0$, the lower bounds on $LATE_{at}$ and $LATE_{nt}$ become zero; otherwise, their upper bounds become zero. Consequently, depending on the sign of $LATE_c$, equation (1) implies that either the lower or upper bound on the ATE equals the ITT effect (which equals $\pi_c LATE_c$ since $\pi_c = p_{1|1} - p_{1|0}$). When the outcome is binary, the bounds in Proposition 2 coincide with those in BSV and Chiburis (2010b), both of which equal the bounds in Shaikh and Vytlacil (2011) and Chesher (2010) when there are no exogenous covariates other than the binary instrument. Moreover, if $LATE_c$ is positive (negative) and Assumptions 1 to 6 hold, then the bounds in Proposition 2 equal the bounds obtained by imposing the mean independence of the IV assumption and the increasing (decreasing) MTR assumption in Manski and Pepper (2000). Importantly, MTR imposes monotonicity of the outcome in the treatment at the individual level, and it requires one to know the direction of the effect a priori. Similarly, depending on the sign of the individual effect, BSV showed the equivalence of their bounds to those under the mean independence of the IV assumption and the MTR assumption for the case of a binary outcome. Thus, in the present setting, our results can be seen as an extension of those in BSV to the case of a non-binary outcome.[5]

[5]See BSV for a discussion of the trade-off between the MTR assumption of Manski and Pepper (2000) and the

2.3 Bounds on the ATE under Weak Mean Dominance across Strata

In practice, some strata tend to have characteristics that make them more likely to have higher mean potential outcomes than others. The three alternative assumptions below formalize the notion that, under the same treatment status, never takers have the highest average potential outcomes among the three strata, while always takers have the lowest. Other rankings across strata, which may be more appropriate for other applications, are certainly possible. The particular direction of the weak mean dominance assumptions we employ is consistent with our analysis of the effectiveness of JC, as we discuss in Section 3.2. We consider three alternative mean dominance assumptions to provide more options to applied researchers wanting to implement our bounds, as some of them may be more plausible than others in certain applications.

Assumption 7a. $E[Y(d)|at] \leq E[Y(d)|nt]$ for $d = 0, 1$.
Assumption 7b. $E[Y(0)|at] \leq E[Y|Z = 0, D = 0]$ and $E[Y(1)|nt] \geq E[Y|Z = 1, D = 1]$.
Assumption 7c. $E[Y(0)|at] \leq E[Y(0)|c]$ and $E[Y(1)|nt] \geq E[Y(1)|c]$.

The always takers and never takers are likely to be the most "extreme" strata in many applications, so Assumption 7a may be viewed as the weakest of the three. Assumption 7b compares the mean $Y(0)$ and $Y(1)$ of the always takers and never takers, respectively, to those of a weighted average of the other two strata, while Assumption 7c compares them to those of the compliers.[6] Note that it is possible for Assumption 7b to hold even if either Assumption 7a or 7c does not hold, providing a middle ground between Assumptions 7a and 7c in some applications. For instance, it is possible to have $E[Y(0)|at] > E[Y(0)|c]$ and $E[Y(0)|at] \leq E[Y|Z = 0, D = 0]$, if $E[Y(0)|nt]$ and the proportions of compliers and never takers are such that the latter inequality holds. Huber et al. (2015) consider an assumption similar in spirit to Assumption 7c, but they do not consider assumptions similar to 7a or 7b (nor Assumption 6).[7] Although none of these assumptions is directly testable, it is possible to obtain indirect evidence about their plausibility by comparing relevant average pre-treatment characteristics—e.g., pre-treatment outcomes—of the different strata (e.g., Flores and Flores-Lagunes, 2010, 2013; Bampasidou et al., 2014; Chen and Flores, 2014). For Assumption 7c, the direction may also be informed by comparing point identified quantities, $E[Y(1)|at]$ to $E[Y(1)|c]$, and $E[Y(0)|nt]$ to $E[Y(0)|c]$, to the extent that the inequalities in Assumption 7c also hold under the alternative treatment status.

We present bounds under Assumptions 1 through 5 and each of the three versions of Assumption 7. Due to the direction of the mean dominance inequalities in Assumption 7, in each case the lower bound is higher than that in Proposition 1, while the upper bound is the same.

assumption of monotonicity of the treatment in the instrument at the individual level.

[6] Note that $E[Y|Z = 0, D = 0] = \frac{\pi_c}{\pi_c+\pi_{nt}} E[Y(0)|c] + \frac{\pi_{nt}}{\pi_c+\pi_{nt}} E[Y(0)|nt]$, with an analogous equation holding for $E[Y|Z = 1, D = 1]$.

[7] They assume the mean potential outcomes of compliers are not lower than those of always and never takers.

Proposition 3 *Let* $UB = \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}$*. Then,*
(a) Under Assumptions 1 through 5 and 7a the bounds $LB \leq ATE \leq UB$ *are sharp, where*

$$LB = \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \overline{Y}^{01} p_{0|1} - \overline{Y}^{10} p_{1|0};$$

(b) Under Assumptions 1 through 5 and 7b the bounds $LB \leq ATE \leq UB$ *are sharp, where*

$$LB = \overline{Y}^{11} - \overline{Y}^{00};$$

(c) Under Assumptions 1 through 5 and 7c the bounds $LB \leq ATE \leq UB$ *are sharp, where*

$$LB = \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}} p_{0|1} - \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}} p_{1|0}.$$

We now consider the combination of Assumption 6 with Assumptions 7a through 7c. In this case, if $LATE_c < 0$, there are testable implications because the following inequalities are expected to hold: $\overline{Y}^{01} \leq \overline{Y}^{10}$ (under Assumption 7a); $\overline{Y}^{01} \leq \overline{Y}^{00}$ and $\overline{Y}^{11} \leq \overline{Y}^{10}$ (under 7b); $\overline{Y}^{01} \leq E[Y(0)|c]$ and $E[Y(1)|c] \leq \overline{Y}^{10}$ (under 7c). If any of these inequalities is rejected in a given application, then the data provide statistical evidence against the validity of the corresponding assumptions. The following three propositions provide the resulting bounds when Assumptions 6 and each one of Assumptions 7a through 7c are combined.

Proposition 4 *Under Assumptions 1 through 6 and 7a the bounds* $LB \leq ATE \leq UB$ *are sharp, where, if* $E[Y|Z=1] - E[Y|Z=0] > 0$*,*

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \max\{\overline{Y}^{10}, \overline{Y}^{01}\} p_{0|1} - \min\{\overline{Y}^{10}, \overline{Y}^{01}\} p_{1|0} \\ UB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$*,*

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \overline{Y}^{01} p_{0|1} - \overline{Y}^{10} p_{1|0} \\ UB &= E[Y|Z=1] - E[Y|Z=0]. \end{aligned}$$

Proposition 5 *Under Assumptions 1 through 6 and 7b the bounds* $LB \leq ATE \leq UB$ *are sharp, where, if* $E[Y|Z=1] - E[Y|Z=0] > 0$*,*

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \max\{\overline{Y}^{10}, \overline{Y}^{11}\} p_{0|1} - \min\{\overline{Y}^{01}, \overline{Y}^{00}\} p_{1|0} \\ UB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$*,*

$$\begin{aligned} LB &= \overline{Y}^{11} - \overline{Y}^{00} \\ UB &= E[Y|Z=1] - E[Y|Z=0]. \end{aligned}$$

Proposition 6 *Under Assumptions 1 through 6 and 7c the bounds $LB \leq ATE \leq UB$ are sharp, where, if $E[Y|Z=1] - E[Y|Z=0] > 0$,*

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \max\{\overline{Y}^{10}, \frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}}\} p_{0|1} \\ &\quad - \min\{\overline{Y}^{01}, \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}}\} p_{1|0} \\ UB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + y^u p_{0|1} - y^l p_{1|0}; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$,

$$\begin{aligned} LB &= \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + \frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}} p_{0|1} - \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}} p_{1|0} \\ UB &= E[Y|Z=1] - E[Y|Z=0]. \end{aligned}$$

Note that, if $LATE_c < 0$, the bounds in Propositions 4 through 6 do not require boundedness of the outcome because Assumption 6 improves upon the upper bound in Proposition 1, while Assumption 7 improves upon the lower bound. In contrast, if $LATE_c > 0$, Assumptions 6 and 7 each improve only upon the lower bound in Proposition 1. Also, the bounds in Propositions 4 through 6 are narrower than the bounds in Proposition 2 and the corresponding bounds in Proposition 3. This is because, under the combined assumptions, the weak monotonicity assumption on the local average outcomes (Assumption 6) improves further upon either the lower or upper bound in Proposition 3, depending on the sign of $LATE_c$, while the weak mean dominance assumptions further improve upon the lower bound in Proposition 2. Hence, relative to the bounds in Huber et al. (2015) that use all their assumptions, the addition of Assumption 6 results in narrower bounds.

The bounds in Proposition 5 coincide with the bounds derived by Chiburis (2010a) under the MTR assumption (without specifying the direction a priori), the decreasing monotone treatment selection or MTS assumption (a special case of the MIV assumption, where the instrument is the realized treatment), and the mean independence of the IV assumption. This is because Assumption 7b coincides with the decreasing MTS assumption imposed on the counterfactual average outcomes of always takers and never takers (i.e., $E[Y(0)|at]$ and $E[Y(1)|nt]$).

As a final note, the bounds in Proposition 6 are also the sharp bounds for the ATE if we replace Assumption 7c with the assumption $E[Y(d)|at] \leq E[Y(d)|c] \leq E[Y(d)|nt]$ for $d = 0, 1$. Interestingly, however, since $E[Y(d)|c]$ may be more difficult to estimate in practice than $E[Y|Z=d, D=d]$ (e.g., if the IV is weak and $p_{1|1} - p_{1|0}$ is close to zero), the estimated bounds in Proposition 5 (using Assumption 7b) could produce narrower confidence intervals in practice than the estimated bounds based on Proposition 6.

2.4 Bounds on the ATT

This subsection presents bounds on the average treatment effect on the treated (ATT). They are derived using the same approach and under similar assumptions to those employed above to derive bounds on the ATE. The treated subpopulation is a mixture of the compliers and always takers strata (e.g., see footnote 4). Hence, we start by defining $q_z \equiv \Pr(Z = z)$ and $r_1 \equiv \Pr(D = 1)$ in order to write the ATT as a weighted average of $LATE_{at}$ and $LATE_c$ as:

$$
\begin{aligned}
ATT &= \frac{q_1}{r_1}(\pi_c LATE_c + \pi_{at} LATE_{at}) + \frac{q_0 \pi_{at}}{r_1} LATE_{at} \qquad (3) \\
&= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - E[Y(0)|at])]. \qquad (4)
\end{aligned}
$$

Equation (4) shows the ATT can also be viewed as a weighted average of the ITT effect and the $LATE_{at}$. Thus, only assumptions on $Y(0)$ for always takers are required to bound the ATT.

The propositions below, labeled Proposition 1' to Proposition 6', present the bounds for the ATT and parallel those presented in the previous subsections for the ATE.

Proposition 1' *Under Assumptions 1 through 5 the bounds $lb \leq ATT \leq ub$ are sharp, where*

$$
\begin{aligned}
lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^u)] \\
ub &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^l)].
\end{aligned}
$$

The next proposition adds the assumption of weak monotonicity of average potential outcomes within the at stratum, which corresponds to Assumption 6 as applied to this stratum (i.e., ignoring the nt stratum). Similar to the case of the ATE, when this assumption is added either the lower bound or the upper bound on the ATT is improved with respect to the benchmark bounds in Proposition 1', depending on the sign of the $LATE_c$.

Proposition 2' *Under Assumptions 1 through 5, and Assumption 6 as applied to the at stratum, the bounds $lb \leq ATT \leq ub$ are sharp, where, if $E[Y|Z=1] - E[Y|Z=0] > 0$,*

$$
\begin{aligned}
lb &= \frac{q_1}{r_1}(E[Y|Z=1] - E[Y|Z=0]) \\
ub &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^l)];
\end{aligned}
$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$,

$$
\begin{aligned}
lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^u)] \\
ub &= \frac{q_1}{r_1}(E[Y|Z=1] - E[Y|Z=0]).
\end{aligned}
$$

For the ATT, the three alternative mean dominance assumptions we consider are the same as those in Assumptions 7a to 7c with respect to the non-identified mean $E[Y(0)|at]$ (i.e., ignoring those inequalities involving $E[Y(1)|nt]$). As for the ATE, only the lower bounds on the ATT are improved under the mean dominance assumption relative to the bounds in Proposition 1'.

Proposition 3' *Let* $ub = \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^l)]$. *Then,*
(a) Under Assumptions 1 through 5, and 7a as applied to the term $E[Y(0)|at]$*, the bounds* $lb \leq ATT \leq ub$ *are sharp, where* $lb = \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \overline{Y}^{10})]$*;*
(b) Under Assumptions 1 through 5, and 7b as applied to the term $E[Y(0)|at]$*, the bounds* $lb \leq ATT \leq ub$ *are sharp, where* $lb = \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \overline{Y}^{00})]$*;*
(c) Under Assumptions 1 through 5, and 7c as applied to the term $E[Y(0)|at]$*, the bounds* $lb \leq ATT \leq ub$ *are sharp, where* $lb = \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \frac{\overline{Y}^{00}p_{0|0} - \overline{Y}^{10}p_{0|1}}{p_{1|1} - p_{1|0}})]$.

The last three propositions provide bounds on the ATT combining Assumption 6 with each of Assumptions 7a to 7c (all these four assumptions as applied only to the non-identified mean $E[Y(0)|at]$). As with the ATE, when the $LATE_c$ is positive, both the weak monotonicity of average potential outcomes assumption (Assumption 6) and the weak mean dominance assumption (Assumptions 7a to 7c) improve the lower bound on the ATT relative to the bounds in Proposition 1'. Under a negative $LATE_c$, Assumption 6 and the different versions of Assumption 7 improve the upper and lower bounds on the ATT, respectively, thus eliminating the requirement that the outcome is bounded. In general, as for the ATE, the weak monotonicity of average potential outcomes assumption for the at stratum has identifying power for the ATT.

Proposition 4' *Under Assumptions 1 through 5, and 6 and 7a as applied to the term* $E[Y(0)|at]$*, the bounds* $lb \leq ATT \leq ub$ *are sharp, where, if* $E[Y|Z=1] - E[Y|Z=0] > 0$*,*

$$\begin{aligned} lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \min\{\overline{Y}^{01}, \overline{Y}^{10}\})] \\ ub &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^l)]; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$*,*

$$\begin{aligned} lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \overline{Y}^{10})] \\ ub &= \frac{q_1}{r_1}(E[Y|Z=1] - E[Y|Z=0]). \end{aligned}$$

Proposition 5' *Under Assumptions 1 through 5, and 6 and 7b as applied to the term* $E[Y(0)|at]$*, the bounds* $lb \leq ATT \leq ub$ *are sharp, where, if* $E[Y|Z=1] - E[Y|Z=0] > 0$*,*

$$\begin{aligned} lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \min\{\overline{Y}^{01}, \overline{Y}^{00}\})] \\ ub &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - y^l)]; \end{aligned}$$

and if $E[Y|Z=1] - E[Y|Z=0] < 0$*,*

$$\begin{aligned} lb &= \frac{1}{r_1}[q_1(E[Y|Z=1] - E[Y|Z=0]) + p_{1|0}(\overline{Y}^{01} - \overline{Y}^{00})] \\ ub &= \frac{q_1}{r_1}(E[Y|Z=1] - E[Y|Z=0]). \end{aligned}$$

Proposition 6' *Under Assumptions 1 through 5, and 6 and 7c as applied to the term $E[Y(0)|at]$, the bounds $lb \leq ATT \leq ub$ are sharp, where, if $E[Y|Z=1]-E[Y|Z=0]>0$,*

$$
\begin{aligned}
lb &= \frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-\min\{\overline{Y}^{01},\frac{\overline{Y}^{00}p_{0|0}-\overline{Y}^{10}p_{0|1}}{p_{1|1}-p_{1|0}}\})] \\
ub &= \frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-y^l)];
\end{aligned}
$$

and if $E[Y|Z=1]-E[Y|Z=0]<0$,

$$
\begin{aligned}
lb &= \frac{1}{r_1}[q_1(E[Y|Z=1]-E[Y|Z=0])+p_{1|0}(\overline{Y}^{01}-\frac{\overline{Y}^{00}p_{0|0}-\overline{Y}^{10}p_{0|1}}{p_{1|1}-p_{1|0}})] \\
ub &= \frac{q_1}{r_1}(E[Y|Z=1]-E[Y|Z=0]).
\end{aligned}
$$

2.5 Estimation and Inference

The objects in the expressions of the bounds derived above can be estimated with sample analogs. However, complications for estimation and inference arise in the bounds that involve minimum (min) or maximum (max) operators. First, because of the concavity (convexity) of the min (max) function, sample analog estimators of the bounds can be severely biased in small samples. Second, closed-form characterization of the asymptotic distribution of estimators for parameters involving min or max functions are very difficult to derive and, thus, usually unavailable. Furthermore, Hirano and Porter (2012) showed that there exist no locally asymptotically unbiased estimators and no regular estimators for parameters that are nonsmooth functionals of the underlying data distribution, such as those involving min or max operators.

To deal with those issues, for bounds containing min or max operators we employ the methodology proposed by Chernozhukov, Lee and Rosen (2013; hereafter CLR) to obtain confidence regions for the true parameter value, as well as half-median unbiased estimators for the lower and upper bounds. The half-median-unbiasedness property means that the upper (lower) bound estimator exceeds (falls below) the true value of the upper (lower) bound with probability at least one half asymptotically. This is an important property because achieving local asymptotic unbiasedness is not possible, implying that bias-correction procedures cannot completely eliminate local bias, and reducing bias too much will eventually cause the variance of the procedure to diverge (Hirano and Porter, 2012). For details on our implementation of CLR's method see Flores and Flores-Lagunes (2013). For the bounds without min or max operators, we use sample analog estimators and construct the confidence regions for the true parameter value proposed by Imbens and Manski (2004).

3 Bounds on Average Treatment Effects of Job Corps Training

3.1 The Job Corps Program and Data

Job Corps (JC) is the largest and most comprehensive education and job training program in the United States. It serves economically disadvantaged youth through the delivery of academic education, vocational training, residential living, health care and health education, counseling, and job placement assistance. Since its creation in 1964, JC has served over 2 million young people (U.S. Department of Labor, 2015). Eligibility into the program is based on age (16 to 24), being economically disadvantaged, being high school dropout or in need of additional education or vocational training, not being on probation or parole; and being free of serious medical or behavioral problems. Approximately 70% of JC enrollees are members of minority groups, and 75% are high school dropouts (U.S. Department of Labor, 2015). The average length of stay for participants is 8.2 months, with an average number of academic and vocational hours received in JC comparable to that of a regular year of high school education (Schochet et al., 2001).

In the mid-1990s, the U.S. Department of Labor funded the National Job Corps Study (NJCS) to assess the program effectiveness. We use data from the NJCS, whose main feature was the randomization of eligible applicants into a treatment group allowed to enroll in JC and a control group barred from receiving JC services for three years. Eligible applicants were taken at random from the 48 contiguous U.S. states, making this social experiment one of the few with nationally representative character. From a randomly selected research sample of 15,386 first time eligible applicants, 61 percent (9,409) were assigned to the treatment group and 39 percent (5,977) to the control group. These individuals were interviewed at baseline (randomization) and followed with surveys at weeks 52, 130, and 208 after randomization (Schochet et al., 2001).

Randomization in the NJCS took place before participants' assignment to a JC center. As a result, there is an important degree of non-compliance as only about 73% of individuals in the treatment group actually enrolled in JC, while about 1.4% of individuals in the control group managed to enroll in JC during the three-year embargo (Schochet et al., 2001). Counting individuals in the control group that enrolled in JC after the embargo was lifted, the latter percentage increases to 4.3%. Non-compliance is a very common occurrence in randomized experiments, which typically forces researchers to change their original goal of estimating the causal effect of receiving treatment for the population (e.g., eligible applicants) or those receiving treatment (e.g., JC participants), to that of estimating effects for a different treatment or subpopulation. For example, in order to take full advantage of randomization, most of the previous evaluations of JC using the NJCS data estimate the ITT effect or the $LATE_c$ (e.g., Burghardt et al., 2001; Schochet et al., 2001; Schochet et al., 2008; Lee, 2009; Flores-Lagunes et al., 2010). In the case of the ITT effect, the randomization indicator is employed in lieu of the actual treatment receipt indicator, which implies that the effect being estimated is that of being offered participation in JC, rather than the effect of actual JC participation. As a result, focusing on ITT effects tends to dilute the impacts of JC (e.g., Schochet et al., 2001; Chen and Flores, 2014). In the case of the $LATE_c$, the randomization indicator is used as an IV for actual program enrollment, which identifies the effect of JC participation for the subpopulation of compliers (individuals who participate in JC only if assigned to enroll). In our application, the results below show this effect is representative of only about 69% of eligible JC applicants.

To our knowledge, the previous literature on the effectiveness of JC using data from the NJCS has not analyzed the effects of JC participation on the population of eligible applicants (ATE) or the group of participants (ATT), both of which are very important populations from a policy perspective. We fill this gap by undertaking inference on these two parameters. The outcome variables we consider are weekly earnings and employment at week 208 after random assignment, and public assistance benefits received during the fourth year after randomization.[8]

To conduct our analysis, we start with the original NJCS sample of individuals that responded to the 48-month interview (11,313 individuals, 4,485 in control and 6,828 in treatment groups) and drop cases with missing information on three key variables: the outcomes, the randomization indicator, and the indicator for actual enrollment in JC. Given that the cases with missing information on labor market outcomes (weekly earnings and employment) and receipt of public benefits are different, we construct two samples. The first sample, for labor market outcomes, consists of 10,520 individuals (4,187 in control and 6,333 in treatment groups), while the second sample, for receipt of public benefits, consists of 10,976 individuals (4,387 in control and 6,589 in treatment groups). In some analyses below we employ pre-treatment variables, which may be missing for some individuals. In those cases, we impute the missing information using the mean of the corresponding variable. Throughout the analysis, we employ NJCS-provided design weights, since due to both design and programmatic reasons some subpopulations had different sampling probabilities (Schochet et al., 2001).[9]

Table 1 reports a selection of average baseline characteristics for both samples by random assignment status (Z), along with the percentage of missing values for each variable. As one would expect given the randomization in the NJCS, and consistent with the original NJCS reports, the differences in average pre-treatment characteristics between treatment and control groups are statistically insignificant in both samples.[10] Thus, both samples maintain the balance of baseline variables between treatment and control groups. The means of the variables are also in line with the characteristics of eligible JC applicants in other studies (e.g., Schochet et al., 2001; Schochet et al., 2008; Lee, 2009; Flores-Lagunes et al., 2010). For instance, the typical individual is 18 years old, a minority, never married, without a job in the previous year, with low weekly earnings (about $110), and received

[8]Public benefits include Aid to Families with Dependent Children (AFDC) or Temporary Assistance for Needy Families (TANF), food stamps, Supplemental Security Income (SSI) or Social Security Retirement, Disability, or Survivor (SSA), and General Assistance.

[9]Specifically, the weights we employ address sample design, 48-month interview design, and 48-month interview non-response.

[10]The exceptions are the differences in means for "personal income between 6,000 and 9,000" in both samples, and "Father's education" (which is marginally statistically significant) in the public benefits sample.

public benefits (59% of eligible applicants did).

3.2 Assessment of Assumptions and Preliminary Estimates

In this subsection we undertake an assessment of our assumptions in the context of evaluating the effects of JC, and also discuss some preliminary estimates of objects that are point identified.

Assumption 1 is random assignment of the instrument, which in our context is satisfied by design. Assumption 2 is the exclusion restriction assumption, which states that random assignment (the instrument) has an effect on the outcomes exclusively through enrollment in JC (the treatment). This assumption is likely satisfied in the present context, and has been widely used in the JC literature (e.g., Schochet et al., 2001; Frumento et al., 2012; Chen and Flores, 2014). However, there could be threats to its validity. For instance, this assumption could be violated if some individuals become overly discouraged by receiving the random control assignment that their labor market outcomes or public benefits are directly affected. As has been argued elsewhere (e.g., Schochet et al., 2001; Frumento et al., 2012), while this type of responses may directly affect the short run outcomes of those individuals, it seems reasonable to assume that assignment to JC has a negligible effect on the long run outcomes we consider through channels other than JC participation.

The top panel of Table 2 shows some relevant estimates for our two analysis samples (each sample in a vertical panel of the table). The first two rows show estimated averages for the groups with $Z = 1$ (treatment group) and $Z = 0$ (control group). By looking at the column "Enrollment" (in JC) in each vertical panel, it is clear that non-compliance behavior is similar between the two samples: 73% of individuals in the treatment group enrolled in JC, while 4.3% of individuals in the control group enrolled in JC at some point during the 208 weeks after randomization. The entries in the other columns show the mean outcomes in each of the groups with $Z = 1$ and $Z = 0$. Assumption 3 states that the instrument has a non-zero average effect on the treatment. This is clearly the case in each sample by looking at the ITT estimates on "Enrollment" (third row). The estimated effect of the instrument on the treatment is a highly statistically significant 0.69 in both samples. The other estimated ITT effects on that row pertain to the outcomes and are highly statistically significant as well. They are \$22.19, 0.038 percentage points, and $-$\$84.29 for weekly earnings, employment (both at week 208 after randomization), and public benefits during year 4 after randomization, respectively.

To point identify the average effect of JC participation for compliers ($LATE_c$), individual-level weak monotonicity of the treatment in the instrument (Assumption 4) is needed (IA; AIR). Although this is a conventional assumption of IV-methods, it is strong in some applications because monotonicity is imposed at the individual level. Not surprisingly, this assumption has considerable identifying power for $LATE_c$ (allowing point identification), and for our bounds on ATE and ATT (see footnote 3). In our context, Assumption 4 requires that no individual enrolls in JC if assigned to the control group but does not enroll if assigned to participate in JC. This assumption has been used previously in the JC literature (e.g., Schochet et al., 2001; Frumento et al., 2012; Chen and Flores, 2014), and seems plausible since it is unlikely that eligible applicants would enroll in JC only if denied access to it. The fourth row in the top panel of Table 2 presents $LATE_c$ estimates under Assumptions 1 through 4. The estimates are \$32.29, 0.055, and $-$\$122.28 for weekly earnings, employment, and public benefits, respectively; and are all highly statistically significant. As usual, the $LATE_c$ estimates are larger in absolute value than the corresponding ITT estimates because the former equal the latter divided by the effect of assignment on enrollment. These results on the ITT and $LATE_c$ are consistent with the findings from the NJCS in Burghardt et al. (2001).

Under Assumptions 1 to 4 it is also possible to point identify the proportion of each stratum in the population. These point estimates are shown in the second panel of Table 2. In both samples, the proportion of compliers is the largest (69%), followed by never takers (27%) and always takers (4%). Hence, 69 percent of the individuals would enroll in JC if offered the opportunity to do so, and would not enroll otherwise. Importantly, about 1 in 4 eligible individuals—the never takers—decides not to participate in JC regardless of whether or not they are offered the opportunity to do so. This is an important subpopulation from a policy perspective, as these individuals are part of the target population of JC but would not participate in the program even if given the opportunity. If their outcomes could be improved by participating in JC, there could be gains from finding ways to encourage them to enroll in JC. Thus, it is important to gather evidence on the average effects of JC for never takers ($LATE_{nt}$), and to consider the characteristics of the individuals in this stratum to shed light on the possible reasons why they decide not to participate in JC. We touch on these points in the empirical analysis below.

The next assumptions to consider are those added to the usual IV assumptions to construct our bounds. Assumption 5 is that the outcome is bounded. Employment is naturally bounded in $[0, 1]$. A common practice in the bounding literature, which we also adopt here, is to use the observed in-sample maximum and minimum values for outcomes such as earnings and public benefits. Assumption 6 imposes weak monotonicity in the treatment of the mean outcomes of each stratum. Since under Assumptions 1 to 4 $LATE_c$ is point identified, Assumption 6 becomes an assumption on the signs of $LATE_{at}$ and $LATE_{nt}$, which under this assumption are identified from the sign of $LATE_c$ for each outcome. Given the $LATE_c$ estimates in Table 2, Assumption 6 imposes non-negative $LATE$s for weekly earnings and employment, and non-positive $LATE$s for public benefits. Based on the characteristics of JC and its stated goals (see Section 3.1), along with the long-term nature of the outcomes we consider (which mitigate potential "lock-in" effects—van Ours, 2004), we expect that, on average, the effects of JC on always and never takers will have the postulated signs in Assumption 6 for each outcome. Put differently, we would not expect that, on average, JC would harm the outcomes of always or never takers. Empirical evidence on the plausibility of Assumption 6 can be gathered by analyzing estimated bounds on $LATE_{nt}$ and $LATE_{at}$ that do not impose this assumption. Below we present cases in which such bounds are able to determine the sign of these parameters in our application, with the results being consistent with the directions implied by Assumption 6.

Assumption 7 imposes weak mean dominance of potential outcomes across different strata. As mentioned in Section 2.3, the (weak) ranking of the average potential outcomes of strata can be informed by the estimates of mean outcomes point identified under Assumptions 1 through 4. The third panel of Table 2 reports a number of estimated mean outcomes for different strata and observed groups, and the last panel shows relevant average outcome differences.[11,12] The estimated means follow a certain pattern in each of the samples: under the treatment status, the mean outcome for always takers ($E[Y(1)|at]$) is the smallest, followed by the mean outcome for the mixture of always takers and compliers ($E[Y|Z=1,D=1]$), and the mean outcome for compliers ($E[Y(1)|c]$). Under the control status, the mean outcome for compliers ($E[Y(0)|c]$) is the smallest, followed by the mean outcome for the mixture of never takers and compliers ($E[Y|Z=0,D=0]$), and the mean outcome for never takers ($E[Y(0)|nt]$). This ordering is consistent with the general notion of Assumption 7—that the compliers have better average potential outcomes than the always takers, but worse than the never takers.

We now employ these point estimates to inform the plausibility of the different versions of Assumption 7. Although it is not possible to compare the mean of the same potential outcomes for all three strata, the estimated mean outcomes in Table 2 suggest that never takers and always takers are the two "extreme" groups pertaining to their mean outcomes. Thus, Assumption 7a that compares these two strata under the same treatment status seems plausible. Regarding Assumption 7b, given that the never takers appear to have the more favorable outcomes, followed by the compliers and then the always takers, it seems plausible in our application. Moreover, for the labor market outcomes under Assumption 6 (so $LATE_{at} \geq 0$), the fact that $\overline{Y}^{00} = E[Y|Z=0,D=0]$ is (statistically) significantly larger than $\overline{Y}^{01} = E[Y(1)|at]$ implies that the first inequality in Assumption 7b ($E[Y|Z=0,D=0] \geq E[Y(0)|at]$) holds.[13] Regarding Assumption 7c, the inequalities $E[Y(1)|at] \leq E[Y(1)|c]$ and $E[Y(0)|c] \leq E[Y(0)|nt]$ can shed light on the plausibility of this assumption to the extent that these relationships also hold under the alternative treatment status. As shown in Table 2, $E[Y(1)|c]$ is statistically greater than $E[Y(1)|at]$ for both labor market outcomes, providing indirect evidence in favor of the first inequality of Assumption 7c ($E[Y(0)|at] \leq E[Y(0)|c]$). Similarly, for weekly earnings, $E[Y(0)|nt]$ is statistically larger than $E[Y(0)|c]$ with a 0.10 significance level, offering indirect evidence in favor of the second inequality of Assumption 7c ($E[Y(1)|nt] \geq E[Y(1)|c]$). The rest of the comparisons are not statistically different from zero, providing no indirect evidence against Assumption 7c. Lastly, recall that for the case

[11]Remember that the non-identified quantities are $E[Y(0)|at]$ and $E[Y(1)|nt]$, while the averages in Table 2 provide estimates for the mean of $Y(0)$ for compliers, never takers, and a mixture of them; and of $Y(1)$ for compliers, always takers, and a mixture of them.

[12]We follow Lee (2009) and use a transformed measure of weekly earnings and public benefits to minimize the effect of outliers in the estimation of sample means. Specifically, the observed outcome distribution for each of those two outcomes is split into 20 percentile groups ($5^{th}, 10^{th}, \ldots, 95^{th}$), and the mean outcome within each of the 20 groups is assigned to each individual.

[13]A similar argument could be made for the second inequality in Assumption 7b if $\overline{Y}^{10} = E[Y(0)|nt]$ were statistically larger than $\overline{Y}^{11} = E[Y|Z=1,D=1]$, but for the labor market outcomes they are not statistically different from each other. Note, however, that this does not contradict the assumptions, as it is still possible to have $E[Y(1)|nt] \geq E[Y(0)|nt]$ (Assumption 6) and $E[Y(1)|nt] \geq E[Y|Z=1,D=1]$ (Assumption 7b).

of public benefits, in which the estimated $LATE_c$ is negative, there are testable implications under Assumptions 1 through 7 (see Section 2.3). They are shown in the last five rows of Table 2, with the first of them corresponding to Assumption 7a, the next two to Assumption 7b, and the last two to Assumption 7c. All five testable implications are soundly satisfied in our application. Overall, the estimated average outcomes in Table 2 do not provide evidence against the different versions of Assumption 7, and their ordering conforms to that implied by Assumption 7.

An additional way to gather indirect evidence on Assumption 7 is to compare average baseline characteristics of the strata that are likely to be highly correlated with the outcomes considered (e.g., Flores and Flores-Lagunes, 2010, 2013; Bampasidou, et al., 2014). For instance, Assumption 7c would be less likely to hold for a particular outcome if the average baseline characteristics likely to be highly correlated to that outcome would make compliers more likely to have higher mean potential outcomes than never takers, or lower mean outcomes than always takers. Similarly, Assumption 7a would be less likely to hold if those average baseline characteristics for the always takers would make them more likely to have higher mean potential outcomes than the never takers.[14] In addition, comparing the average baseline characteristics of the different strata can help to gain intuition on the results from Table 2 that suggest that never takers may have the highest average potential outcomes for the labor market outcomes (weekly earnings and employment) but also for the public benefits outcome (in both cases followed by compliers and then always takers), which at first may seem counterintuitive.

Tables 3 and 4 (for each analysis sample, respectively) show estimated averages of selected pre-treatment characteristics by stratum, along with differences in averages across strata.[15] The estimates are similar in the two samples. We start by considering the labor market outcomes. The stratum of never takers appears to have average pre-treatment characteristics that are highly related to better labor market outcomes, as individuals in this stratum are more likely to be older, have higher level of education at baseline, have personal income above \$9,000 (and less likely to have personal income below \$3,000), and, importantly, to have better labor market outcomes the year prior to randomization and at baseline (e.g., earnings). By contrast, individuals in the always takers stratum are more likely to be younger, have lower level of education at baseline, lower personal income, and lower earnings in the year prior to randomization—all characteristics that are arguably highly related to worse labor market outcomes. Moreover, looking at the differences in average pre-treatment characteristics between these two strata (last column in each table), all the differences documented above have the expected sign (according to Assumption 7a) and most of them are statistically

[14] For Assumption 7b, one can also compute the average baseline characteristics of the groups $\{Z = 0, D = 0\}$ and $\{Z = 1, D = 1\}$. While we omit these results for brevity, the results shown below for the different strata are also informative of Assumption 7b.

[15] Under Assumptions 1 and 4 the average baseline characteristics of all strata are point identified from the observed mean of those characteristics for the four groups given by the values of $\{Z, D\}$, as each of them is a weighted average of the mean characteristics of different strata (see footnotes 4 and 6), with the weights being point identified. We employ a GMM approach to estimate the average baseline characteristics of the strata because the number of moment conditions exceeds the number of parameters. See Appendix for details.

significant. Particularly notable are the statistical significance of pre-treatment outcomes such as the earnings measures, which are expected to be highly correlated with the labor market outcomes at week 208 after randomization. This indirect evidence favors the plausibility of Assumption 7a for the labor market outcomes.

We now turn our attention to the public benefits outcome. Interestingly, as compared to always takers, never takers also have average baseline characteristics that would make them more likely to receive higher levels of public benefits. In particular, relative to always takers, never takers are more likely to be female, have children, be married, and have household income below $3,000 (and less likely to have household income above $18,000), with all these differences being statistically significant. Similarly, never takers are more likely to receive public benefits at baseline, and to have received them for more months, although these differences are not statistically significant.[16] It is known that the variables previously mentioned are highly correlated to the receipt of public assistance (e.g., Moffitt, 2003). For instance, AFDC/TANF benefits are specifically directed towards families with children. Likewise, the outcome variable public benefits received includes assistance that the individuals, their spouse, or children who lived with them received; hence, individuals who are married with children are likely to receive higher public benefits than single individuals without children. Moreover, Schochet et al. (2001) report that females with children had very different experiences with public benefits—both at baseline and post-randomization—than males and females without children. They indicate that, while 51 percent of males and 67 percent of females received public benefits the year prior to randomization, 88 percent of females with children did. They also report that after randomization females with children continued to receive public benefits with a much larger proportion than males and females without children did, and that their average amount received was also the largest among the three groups (by a considerable amount). In sum, the average of baseline characteristics highly correlated to the receipt of public benefits for the never takers and always takers provide indirect evidence in favor of Assumption 7a for the public benefits outcome.

In addition, the average baseline characteristics of the strata previously discussed shed light on the results in Table 2 that suggest never takers may have the largest average potential outcomes for both the labor market and public benefits outcomes, which is also the intuition behind Assumption 7. The results in Tables 3 and 4 suggest that, while never takers are an homogeneous group with respect to compliance behavior, they are a heterogeneous group in other regards: as compared to the other strata, this stratum is comprised of individuals who, at baseline, are on average better educated and have better labor market histories, but also of individuals who are more likely to be female, married, and have children. As a result, as compared to the other strata, never takers may indeed have higher average potential outcomes for both labor market and public benefits outcomes.

Turning attention to the stratum of compliers, their estimated averages for the pre-treatment characteristics previously discussed are generally in-between the magnitude of the corresponding ones for never takers and always takers. This is consistent with the proposed strata ordering implied

[16]Unfortunately, there is no information on the dollar amount of public benefits received prior to randomization.

in Assumptions 7b and 7c. The fourth and fifth columns in Tables 3 and 4 show estimated differences in average pre-treatment characteristics between compliers and never takers, and compliers and always takers, respectively. Most of the differences between never takers and compliers are of the expected sign for the corresponding outcomes, and are often statistically significant; for example, those regarding education and earnings at baseline for the labor market outcomes, and female and having children for public benefits. While the differences between compliers and always takers are sometimes of the opposite sign to the one conjectured in Assumptions 7b and 7c, in no instance are those opposite-signed differences statistically significant. Importantly, for the labor market outcomes, the differences in earnings in the year prior to randomization and education at baseline are of the expected sign and statistically different; while for the public benefits outcome the female and household income variables are of the expected sign and statistically different. In sum, based on the indirect evidence from average pre-treatment characteristics, we conclude that the data do not provide indirect evidence against the different versions of Assumptions 7, and that the majority of the evidence suggests that these assumptions are plausible for all the outcomes considered.

Before concluding this subsection, we note that the average characteristics of the different strata can provide relevant information to policy makers and JC administrators (e.g., Frumento et al., 2012; Bampasidou et al., 2014). Of particular interest, as mentioned above, is the never takers stratum as these individuals always decide against enrolling in JC regardless of treatment assignment (even though they initially applied to join JC). As noted above, relative to the other strata, never takers tend to have on average better education and labor market histories at baseline, while at the same time they are more likely to be female, married, and have children. Hence, as discussed by Frumento et al. (2012), one possible reason why these individuals decide against enrolling in JC may be that they believe they would not benefit from it (they may consider themselves to be "too good" for the program), while another reason may be that some of them are not able to enroll because of family constraints (e.g., difficulty in finding childcare).[17] JC administrators could use this information to increase the participation of these individuals in JC. A first critical step, however, is to determine whether or not never takers would indeed benefit from JC. If they would benefit, then administrators could, for example, focus their efforts on better informing these individuals that, despite their good characteristics relative to other eligible applicants, they can still benefit from enrolling in JC. In addition, JC administrators could focus on relaxing some of the family constraints that may prevent these individuals from participating in JC, for example, by extending JC's childcare services (which are available at some centers). On the other hand, if these individuals would not benefit from JC, administrators could try to find better ways to serve them (e.g., through alternative services or education programs). In Section 3.5 we analyze the effects of JC for this important stratum.

[17] Frumento et al. (2012) analyze the effect of JC on wages using the NJCS and a "principal stratification" approach to adjust for non-compliance and sample selection (as wages are observed only for employed individuals). Our stratification is different from theirs because they have to address two identification issues, and they rule out the existence of always takers (for further discussion, see also Chen and Flores, 2014). However, in general, our results regarding the characteristics of never takers are consistent with theirs.

3.3 Results on the Bounds on the ATE

Table 5 shows estimated bounds on the ATEs on labor market outcomes (weekly earnings and employment) at week 208 after randomization and on public welfare benefits received during the fourth year after randomization. The vertical panels correspond to each of these outcomes. The ATE is interpreted as the average effect of JC participation for the population of eligible applicants (the target population in the NJCS). Estimated bounds are presented under Assumptions 1 to 4 plus the additional assumptions corresponding to Propositions 1 through 6. Under each pair of estimated bounds in Table 5, we report a 95 percent confidence interval for the true value of the parameter (ATE).

We begin by discussing the estimated bounds on the ATE for weekly earnings. The estimated bounds under Proposition 1 (shown in the first row) represent a benchmark for subsequent bounds. They use the IV assumptions in AIR (Assumptions 1 to 4) plus the bounded-outcome assumption (Assumption 5). The estimated bounds are wide and fail to identify the sign of the ATE. Thus, it is desirable to consider additional plausible assumptions to tighten them. Recall that these bounds coincide with the IV bounds proposed by Manski (1990), Heckman and Vytlacil (2000), and Kitagawa (2009) when applied to our setting. The estimated bounds on the ATE under Proposition 2, which use the additional assumption of weak monotonicity in D of average outcomes of strata (Assumption 6), are presented in the second row. For weekly earnings, the ATE is bounded within the interval [\$22.19, \$201.02], and its corresponding 95 percent confidence interval excludes zero. Thus, the bounds obtained by assuming non-negative $LATE$s for always and never takers imply strictly positive average effects of JC on weekly earnings for eligible applicants. Relative to the estimated bounds under Proposition 1, adding Assumption 6 increases the lower bound to \$22.19, which equals the value of the ITT. Recall that the bounds under Proposition 2 are equivalent to those under the MTR assumption in Manski and Pepper (2000), with the important distinction that Assumption 6 is imposed at the stratum level rather than at the individual level, making it easier to hold in practice.

Rows 3 to 5 in Table 5 present the estimated bounds under each of the three weak mean dominance assumptions across strata (Assumptions 7a to 7c, corresponding to Propositions 3a to 3c). Each one of these assumptions improve upon the lower bound in Proposition 1. Importantly, Assumptions 7a to 7c do not impose restrictions on the signs of the $LATE$s (and thus ATE). The estimated bounds under Proposition 3a are not able to identify the sign of the ATE on weekly earnings. However, in each set of estimated bounds under Propositions 3b and 3c, negative ATEs on weekly earnings are statistically ruled out with 95 percent confidence. Therefore, we are able to pin down the sign of the average effect of JC on weekly earnings for eligible applicants without imposing restrictions on the sign of this effect (since Assumption 6 is not used). Moreover, the different identifying power of Assumptions 7a to 7c is evident in this application—while adding Assumption 7a is not enough to identify the sign of the ATE, Assumptions 7b and 7c are. In this case, Assumption 7c yields the tighter estimated bounds, which are also tighter than the estimated bounds under Proposition 2.

The last three rows in Table 5 show estimated bounds when combining Assumptions 1 to 6 with each one of Assumptions 7a to 7c (Propositions 4 to 6, respectively). Given the positive ITT effect on weekly earnings, in each of the bounds only the lower bound is improved relative to the benchmark bounds in Proposition 1.[18] Each of the three sets of estimated bounds—and the corresponding 95 percent confidence intervals—identify the sign of the ATE on weekly earnings. The estimated bounds in Proposition 6 are the narrowest, [\$24.61, \$201.04]. They imply that the percentage increase in average weekly earnings from participating in JC for eligible applicants is bounded between 11.6% and 94.4%.[19] For these estimated bounds, the lower bound is 10 percent higher than the ITT effect (\$22.19), while $LATE_c$ (\$32.29) falls within the bounds, with both estimated effects falling inside the 95 percent confidence interval for the ATE. However, note that our estimated bounds are for the average effect of actually enrolling in JC (as opposed to the effect of being allowed to enroll in JC—the ITT) for all eligible applicants (as opposed to being only for compliers—the $LATE_c$). Lastly, ATEs of JC on weekly earnings that are lower than \$16.01 (7.5%) and larger than \$210.59 (98.9%) can be ruled out with 95 percent confidence.

The second vertical panel in Table 5 presents the estimated bounds for employment. In contrast to weekly earnings, employment is binary and thus bounded in $[0, 1]$. A similar pattern to the estimated bounds for weekly earnings is found in the bounds for employment. The estimated benchmark bounds under Proposition 1 are $[-0.15,\ 0.163]$, which are wide and unable to identify the sign of the ATE on employment. In the binary-outcome setting, these bounds also coincide with those in Balke and Pearl (1997). When adding the assumption of weak monotonicity in D of average outcomes of strata (Assumption 6), the estimated bounds (and corresponding 95 percent confidence intervals) in Proposition 2 identify the sign of the ATE on employment: [0.038, 0.163]. These bounds are also equal to those proposed by BSV, Chesher (2010), Chiburis (2010b), and Shaikh and Vytlacil (2011), all of whom analyze a binary outcome.

Replacing Assumption 6 with each of Assumptions 7a to 7c (rows 3 to 5) produces estimated bounds that generate a pattern similar to that of weekly earnings. Specifically, Assumption 7a by itself is not enough to identify the sign of the ATE, while Assumptions 7b and 7c are. Thus, as for weekly earnings, we are able to statistically rule out a negative or zero average effect of JC on the probability of employment four years after randomization for eligible applicants without making assumptions about the sign of this effect. Turning to the estimated bounds when combining all assumptions in the last three rows, Proposition 6 employing Assumption 7c (last row) provides the tightest bounds on the ATE for employment: [0.042, 0.163]. As percentage increases with respect to $E[Y|D = 0] = 0.582$, these estimated bounds are [7.2%, 28%]. As for weekly earnings, while the lower bound is 10 percent higher than the ITT effect (0.038), both the ITT effect and $LATE_c$ (0.055) fall within the 95 percent confidence interval corresponding to these bounds. Lastly, with 95

[18] Note that sometimes the upper bound changes very slightly. This is the result of the application of the CLR (2013) procedure to compute half-median unbiased estimates and valid confidence intervals.

[19] These percentages are calculated using $E[Y|D = 0] = 212.98$, since $E[Y(0)]$ is not point identified.

percent confidence, we can rule out ATEs of JC on employment that are lower than 0.023 (4%) and larger than 0.18 (30.9%).

The final vertical panel in Table 5 reports the estimated bounds on the ATE for dependence on public benefits. Besides its public policy relevance, this outcome is important from an illustrative point of view because the ITT effect of assignment to JC on public benefits dependence is negative, and thus the bounds under Propositions 4, 5, and 6 do not require the bounded-outcome assumption. The estimated benchmark bounds under Proposition 1 are wide and largely uninformative. When Assumption 6 is added, the estimated bounds are $[-\$632.86, -\$84.29]$, identifying the sign of the ATE and providing much narrower bounds relative to the benchmark bounds. In this case, Assumption 6 imposes non-positive average effects of JC on public benefit dependence for always takers and never takers, which is informed by the point identified negative $LATE_c$ under the current assumptions. Rows 3 to 5 present the estimated bounds under each one of the Assumptions 7a to 7c. Note that, for public benefits, Assumption 6 has stronger identification power than the mean dominance assumptions. In particular, none of the estimated bounds under Assumptions 7a to 7c allow us to identify the sign of the ATE. However, note that these assumptions do have identifying power as they substantially improve the lower bound relative to the one in Proposition 1. For instance, the estimated lower bound under Assumption 7c is $-\$142.76$, ruling out ATEs of JC on public benefits received below -16.8% (using $E[Y|D = 0] = \$852.12$ as reference point).

Lastly, the last three rows report estimated bounds under Propositions 4 to 6, respectively. For public benefits, both the upper and lower bounds are improved (relative to those in Proposition 1) when considering the combination of Assumption 6 and each of Assumptions 7a to 7c, with the former assumption improving the upper bound and the latter ones improving the lower bound. Importantly, this results in the bounded-outcome assumption being not necessary to derive bounds on the ATE for this outcome. The width of the estimated bounds under Propositions 4 to 6 shrinks considerably relative to the estimated bounds under the previous Propositions. The estimated bounds on the ATE for public benefits under the combined assumptions identify the sign of the ATE. The estimated bounds under Proposition 6—employing Assumption 7c—are the narrowest at $[-\$142.76, -\$84.29]$, implying bounds in percentage terms (relative to $E[Y|D = 0]$) of $[-16.8\%, -9.9\%]$. For this outcome, the upper bound equals the estimated ITT effect, and the $LATE_c$ ($-\$122.28$) falls within the bounds. Finally, with 95 percent confidence we are able to rule out ATEs of JC on public benefits received below $-\$210.62$ (-24.7%) or above $-\$22.13$ (-2.6%).

To summarize, we find statistically positive average effects of JC on weekly earnings and employment four years after randomization for the population of eligible applicants, and statistically negative average effects on the yearly amount of public benefits received.

3.4 Results on the Bounds on the ATT

Table 6 presents estimated bounds on the ATTs—interpreted as the average effects of JC for its participants—on the labor market outcomes and the amount of public benefits received. While, in general, the estimated bounds on the ATT show similar patterns to those for the estimated bounds on the ATE, the former are considerably more informative. This is because, to bound the ATT, we only need to bound the $LATE_{at}$ (see equation (4)) and, in our JC application, always takers account for a relatively small proportion of JC participants ($0.043/0.385 = 11.2\%$, where $Pr(D = 1) = 0.385$). In contrast, for the ATE we need to bound both $LATE_{at}$ and $LATE_{nt}$, where always and never takers account for 31% of the underlying ATE population.

The estimated bounds on the ATT for weekly earnings are shown in the first vertical panel of Table 6.[20] Under the bounded-outcome assumption (Assumption 5), the estimated bounds fail to identify the sign of the ATT. However, the width of the estimated identification region is significantly narrower than that of the corresponding ATE bounds in Table 5. All remaining estimated bounds for weekly earnings in Table 6 identify a statistically positive sign for the ATT. Under Assumption 6 (Proposition 2'), the ATT is bounded between [\$28.67, \$43.47], where Assumption 6 imposes a non-negative $LATE_{at}$ (informed by the positive $LATE_c$ estimate for weekly earnings). The estimated bounds for the ATT under each of the mean dominance assumptions (Assumptions 7a to 7c) in rows 3 to 5 improve upon the lower bound in Proposition 1' by a smaller amount relative to the estimated bounds under Assumption 6. Importantly, however, all three estimated bounds under Assumptions 7a to 7c are able to (statistically) pin down the sign of the ATT for weekly earnings without imposing restrictions on the sign of $LATE_{at}$. Interestingly, in this case the estimated bounds on the ATT combining Assumptions 6 and each of 7a to 7c (rows 6 to 8) do not improve upon the estimated bounds using Assumption 6 only (row 2). This is a finite sample result in that the implementation of the CLR procedure to obtain half-median unbiased estimates yields the former estimated bounds slightly wider than the estimated bounds under Assumption 6 only (which do not contain min or max operators). Consequently, the narrowest estimated bounds for the average effect of JC participation on weekly earnings for JC participants are [\$28.67, \$43.47]. Using the weekly earnings of non-participants as reference (since $E[Y(0)|D = 1]$ is not point identified), these bounds imply that the percentage increase in average weekly earnings is between 13.5% and 20.4%. Moreover, with 95 percent confidence, we can rule out ATTs of JC on weekly earnings lower than \$18.32 (8.6%) and larger than \$54.03 (25.4%).

A similar pattern to that of the estimated bounds on the ATT for weekly earnings is found for employment. Excluding the estimated benchmark bounds (first row), all the other estimated bounds identify a statistically positive ATT on employment—including those that do not impose restrictions on the sign of $LATE_{at}$. The narrowest estimated bounds are obtained under Assumption 6 (second row), [0.049, 0.093], implying bounds on the percentage effects (using $E[Y|D = 0]$ as reference) of [8.4%, 16%]. These bounds allow us to discard ATTs on employment lower than 0.027 (4.5%) and larger than 0.116 (19.9%) with 95 percent confidence.

[20] For reference, the estimated values of $Pr(D = 1)$ and $Pr(Z = 1)$—used to estimate the ATT bounds (see equations (3) and (4))—are (standard errors in parenthesis): 0.385 (0.005) and 0.498 (0.005), respectively, for the labor market outcomes sample, and 0.385 (0.005) and 0.496 (0.005), respectively, for the public benefits sample.

The estimated benchmark bounds on the ATT for public benefits received identify a negative sign for the ATT, which is in stark contrast to the results for the ATE, although the 95 percent confidence interval includes positive values. All the other estimated bounds for public benefits are narrower than the benchmark bounds, as expected. Like for the ATE for public benefits, Assumptions 7a to 7c improve only the lower bound relative to the benchmark bounds, while Assumption 6 improves upon the upper bound (since in this case $LATE_c < 0$). Contrary to the estimated ATT bounds for the labor market outcomes, the estimated bounds under Assumption 6 (imposing $LATE_{at} \leq 0$) are wider relative to all other estimated bounds employing any of the versions of Assumption 7. However, since the upper bound is substantially improved by Assumption 6, these bounds are able to statistically rule out a positive or zero ATT of JC on the receipt of public benefits for JC participants with 95 percent confidence. Combining Assumptions 6 with any of Assumptions 7a to 7c (rows 5 to 8) results in narrower estimated bounds relative to those employing just one of those two assumptions. In particular, the estimated bounds combining Assumptions 6 and 7c are the narrowest at [−\$140.29, −\$108.72] (or, relative to $E[Y|D = 0]$, [−16.5%, −12.8%]), with their corresponding 95 percent confidence interval discarding ATTs below −\$237.89 (−27.9%) and above −\$20.18 (−2.4%).

In sum, we find statistically positive average effects of JC on weekly earnings and employment four years after randomization for JC participants—even without imposing restrictions on the average effect of JC on these outcomes for always takers—as well as statistically negative effects for the amount of public benefits received during the fourth year after randomization.

3.5 Results on Bounds on other Average Treatment Effects

To close Section 3, we discuss results for bounds on other average effects of interest, $LATE_{nt}$ and $LATE_{at}$. As discussed in Section 2, the bounds on these parameters are the building blocks for our bounds on the ATE and ATT. The formulas for the bounds on $LATE_{nt}$ and $LATE_{at}$ corresponding to the assumptions in Propositions 1 to 6 are shown in the Appendix in the analogous Propositions 1" to 6". The estimated bounds and corresponding 95 percent confidence intervals for the three outcomes are shown in Table 7. For brevity, we focus here on the estimated bounds for the average effects of JC participation for never takers ($LATE_{nt}$), as this stratum accounts for 27% of the population and is relevant from a policy perspective.[21] As previously discussed, the individuals in this stratum are part of the target population of JC but decide against participating in the program, even if offered the opportunity to enroll. From a policy perspective, it is thus important to determine if, on average, these individuals would benefit from participating in JC, or if they are making the correct decision of not enrolling in JC because they would not benefit from it.

Table 7 shows that, under Assumptions 1 to 4 and the mean dominance assumption in 7c (Proposition 3"c), there is a statistically positive average effect of JC participation on weekly earnings four

[21] In addition, as can be seen from Table 7, the estimated bounds on $LATE_{at}$ for all three of the outcomes considered are not as informative as those on $LATE_{nt}$, although they indeed provide valuable information (e.g., by ruling out large but plausible effects for always takers).

years after randomization for never takers, with estimated bounds equal to [\$13.03, \$641.87]. Given the point estimate for $E[Y(0)|nt]$ of \$223.79 in Table 2, these bounds imply that JC participation increases the average weekly earnings of never takers by at least 5.8%. Importantly, these results are found without imposing restrictions on the sign of this effect. The estimated bounds based on Propositions 3"b (under Assumption 7b), 5", and 6" (the last two using Assumption 6) also rule out negative and zero values of $LATE_{nt}$, although their corresponding 95 percent confidence intervals include zero. However, the 90 percent confidence interval for $LATE_{nt}$ corresponding to Proposition 6", [1.36, 649.5], excludes zero.[22]

The estimated bounds on $LATE_{nt}$ for employment follow a similar pattern to that for weekly earnings, with the estimated bounds under Propositions 3"b, 3"c, 5", and 6" pointing to positive effects of JC on employment four years after randomization for never takers. While none of the corresponding 95 percent confidence intervals exclude zero (contrary to the case for weekly earnings), the 90 percent confidence intervals based on Propositions 3"b and 6" do ([0.006, 0.416] and [0.001, 0.416], respectively), thus providing (marginal) statistical evidence in favor of a positive $LATE_{nt}$ for employment. The narrowest bounds on $LATE_{nt}$ in this case are [0.025, 0.4] under Proposition 3"c (without imposing restrictions on the sign of $LATE_{nt}$), implying that JC participation increases the probability of employment for never takers by at least 4.2%.

Regarding the receipt of public benefits, the bounds on $LATE_{nt}$ are not able to pin down the sign of this effect. However, some of the estimated lower bounds are informative, with the largest of them ruling out decreases greater than \$172.85 (19.6%) for never takers.

In sum, we find statistically positive average effects of JC participation on labor market outcomes four years after randomization for never takers, even when employing bounds that do not impose restrictions on the sign of these effects. These findings provide statistical evidence in favor of Assumption 6 for never takers (i.e., that $LATE_{nt} \geq 0$) for weekly earnings and employment. In addition, our results imply that never takers could indeed benefit from participating in JC, at least with respect to their labor market outcomes. Recall that, as discussed in Section 3.2, the never takers stratum is comprised of individuals who, as compared to those in the other two strata, are on average more educated and have better labor market histories at baseline, and also of individuals who are more likely to be women, be married, and have children. Therefore, some of these individuals may fail to enroll in JC because they miscalculate the potential benefits of JC or think they are "too good for the program" (probably due to having incomplete information), while others may fail to enroll because—even if they want to enroll—they face constraints related to their family situation that prevent them from participating in JC (e.g., lack of access to childcare). In both cases, it seems that JC administrators could encourage the enrollment of these individuals—who would likely benefit from JC and are already part of its target population—by, for example, providing more information about the services offered within JC and the expected benefits of the program even for individuals

[22] As before, the estimated bounds under Proposition 3"c are narrower than those under Proposition 6" (which adds Assumption 6) because the latter bounds contain *min* or *max* operators and thus employ the CLR procedure.

with relatively good labor market histories and education (as compared to other eligible applicants), or by making it more accessible for individuals with children to enroll in JC (e.g., by expanding the childcare services available at some JC centers, or providing subsidies for childcare).

4 Conclusion

This paper derived sharp nonparametric bounds on the population average treatment effect (ATE) and the average treatment effect on the treated (ATT) within an IV framework, and employed them to evaluate the effectiveness of the Job Corps training program. The bounds, derived by extending the work of Imbens and Angrist (1994) and Angrist et al. (1996), improve upon the benchmark bounds—those using the standard IV assumptions plus a bounded-outcome assumption—and other bounds available in the literature. We introduced two sets of assumptions. The first is monotonicity in the treatment of the average outcomes of never takers and always takers, which is novel to the literature on partial identification of the ATE and ATT. It improves upon similar assumptions that are more difficult to justify in practice because of imposing said monotonicity at the individual level (e.g., Manski and Pepper, 2000). The second set of assumptions imposes mean dominance on potential outcomes across strata. We proposed three such mean dominance assumptions, some of which appear to be new to the literature. An important feature of our bounds is that some of them do not require an outcome with bounded support, which is an assumption typically invoked in practice.

The proposed bounds are used to analyze the average effects of Job Corps (JC) for its eligible applicants (ATE) and its participants (ATT). In addition to being a substantive topic, this application of the proposed bounds helps to illustrate the informational content of the different assumptions considered. JC was evaluated during the mid-nineties through a large-scale, nationally representative social randomized experiment. However, due to extensive non-compliance, estimates of the program effectiveness to date concentrate on intention-to-treat (ITT) effects, or local average treatment effects for the compliers subpopulation ($LATE_c$). Thus, we provide new inference on average effects of actual participation in this important program for other policy-relevant populations under relatively weak assumptions, concentrating on three outcomes: weekly earnings, employment, and public benefits dependency.

Our preferred estimated bounds on the ATE indicate that JC increases weekly earnings of eligible applicants (the target population) by at least \$24.61 (about 11.6%) and employment by at least 4.2 percentage points (about 7.2%), both measured at week 208 after randomization. Importantly, we are able to find statistically positive ATEs of JC on these two labor market outcomes without imposing restrictions on the signs of these average effects for never takers and always takers. We also find that JC decreases dependence on public welfare benefits by at least \$84.29 (about 9.9%), and by no more than \$142.76 (about 16.8%), during the fourth year after randomization. Our preferred estimated bounds on the ATT are much narrower than those on the ATE. They indicate that the average effect of JC for program participants is to increase weekly earnings and employment by at least \$28.67 (about 13.5%) and 4.9 percentage points (about 8.4%), respectively, and at most by \$43.47 (about 20.4%) and 9.3 percentage points (about 16%), respectively. Similar to the case of the ATE, we are able to statistically rule out negative and zero ATTs for these two outcomes without imposing assumptions on the sign of the average effect for always takers. In addition, we find that JC decreases the average amount of public benefits received by JC participants by at least \$108.72 (about 12.8%) and at most \$140.29 (about 16.5%). When comparing these results to the corresponding ITT and $LATE_c$ estimates, in all cases these two estimated effects fall within the corresponding 95 percent confidence intervals from our estimated bounds. Importantly, however, our results apply to the effects of actual JC participation (contrary to the ITT) for all eligible applicants or participants (contrary to the $LATE_c$), and thus are also relevant for policy purposes.

Lastly, as a by-product of our main analysis, we consider the average effect of JC for never takers. This is a key subpopulation for policy purposes because it is comprised of individuals who are part of the target population of JC and who, even if offered the opportunity to enroll, decide not to participate in JC. Slightly more than one out of every four individuals in our sample belongs to this stratum. From a policy perspective, it is thus important to determine whether or not, on average, these individuals would benefit from participating in JC. We find statistical evidence that, indeed, the average labor market outcomes of these individuals would be improved by participating in JC. Employing bounds that do not impose restrictions on the sign of this effect, we find that their average weekly earnings and probability of employment would be improved by at least \$13.03 (5.8%) and 2.5 percentage points (4.2%), respectively. Therefore, it may be in the interest of JC administrators to find better ways to encourage those individuals to participate in JC, such as providing more information about its benefits, or relaxing some of the constraints they may face to enroll in JC (e.g., lack of childcare).

Beyond our JC analysis, this paper illustrates the usefulness of the proposed bounds in making inferences about effects that are not point identified with an IV. Clearly, the approach and methods used herein are not restricted to the important problem of addressing non-compliance in randomized social experiments, as they can also be applied to similar settings where an IV is used to address other identification issues, such as endogeneity. For example, in the natural experiment setting of using the Vietnam-era draft lottery status as an IV to analyze the effect of military service on mortality (AIR, 1996), IV methods identify this effect for individuals who were induced by the draft to serve in the military. In this case, the proposed approach could be used to make inferences on this average effect for the population (ATE) or for those who served in the military (ATT). Finally, this paper also illustrates the insights that can be gained by analyzing the average baseline characteristics and effects of the different strata.

Table 1: Summary Statistics of Baseline Variables

	Labor Market Outcomes Sample				Public Benefits Sample			
	Missing Prop.	$Z = 1$	$Z = 0$	Diff.(Std.Err.)	Missing Prop.	$Z = 1$	$Z = 0$	Diff.(Std.Err.)
Female	0	.417	.407	.009 (.010)	0	.415	.406	.009 (.010)
Age at Baseline	0	18.42	18.38	.035 (.042)	0	18.41	18.38	.031 (.041)
White, Non-hispanic	0	.273	.266	.007 (.009)	0	.274	.269	.005 (.009)
Black, Non-Hispanic	0	.483	.478	.005 (.010)	0	.477	.474	.003 (.010)
Hispanic	0	.171	.179	-.008 (.008)	0	.172	.180	-.008 (.007)
Other Race/Ethnicity	0	.073	.078	-.005 (.005)	0	.076	.076	.000 (.005)
Never Married	.017	.916	.915	.001 (.006)	.020	.914	.915	-.001 (.005)
Married	.017	.020	.022	-.002 (.003)	.020	.020	.022	-.001 (.003)
Living Together	.017	.040	.041	-.001 (.004)	.020	.040	.041	-.001 (.004)
Separated	.017	.024	.022	.002 (.003)	.020	.025	.022	.003 (.003)
Has Child	.007	.181	.184	-.003 (.008)	.008	.181	.183	-.002 (.008)
Number of Children	.011	.253	.248	.005 (.012)	.012	.251	.247	.004 (.012)
Personal Education	.018	10.08	10.09	-.008 (.031)	.021	10.08	10.10	-.019 (.030)
Mother's Education	.194	11.50	11.51	-.011 (.058)	.197	11.49	11.53	-.042 (.057)
Father's Education	.391	11.43	11.54	-.110 (.073)	.394	11.45	11.57	-.127* (.072)
Ever Arrested	.017	.258	.263	-.005 (.009)	.019	.259	.266	-.007 (.009)
Household Inc.: <3000	.368	.252	.258	-.006 (.011)	.371	.250	.255	-.005 (.011)
3000-6000	.368	.201	.204	-.004 (.010)	.371	.198	.208	-.010 (.010)
6000-9000	.368	.116	.111	.006 (.008)	.371	.117	.109	.008 (.008)
9000-18000	.368	.245	.243	.001 (.011)	.371	.246	.241	.005 (.011)
>18000	.368	.187	.183	.003 (.010)	.371	.189	.187	.002 (.010)
Personal Inc.: <3000	.083	.786	.790	-.004 (.008)	.086	.783	.788	-.006 (.008)
3000-6000	.083	.129	.129	.000 (.007)	.086	.130	.131	-.000 (.007)
6000-9000	.083	.055	.046	.009** (.005)	.086	.056	.046	.010** (.004)
>9000	.083	.031	.036	-.005 (.004)	.086	.031	.035	-.004 (.004)
Have Job	.031	.216	.209	.007 (.008)	.034	.219	.211	.009 (.008)
Weekly Hours Worked	0	21.69	21.13	.563 (.417)	0	21.71	21.14	.576 (.407)
Weekly Earnings	0	110.35	104.29	6.059 (4.482)	0	110.66	104.53	6.136 (4.328)
Had Job, Prev. Yr.	.016	.651	.643	.008 (.010)	.019	.653	.646	.007 (.009)
Months Employed,Prev.Yr.	0	3.575	3.516	.058 (.085)	0	3.582	3.518	.064 (.083)
Earnings, Prev.Yr.	.081	2991.8	2873.1	118.65 (109.10)	.084	3020.7	2893.8	126.84 (107.01)
Received Public Benefits	.115	.590	.595	-.005 (.010)	.118	.582	.590	-.008 (.010)
Months Received Benefits	.127	6.554	6.542	.012 (.125)	.129	6.469	6.493	-.024 (.122)
Numbers of Observations	10520	6333	4187		10976	6589	4387	

Note: Z denotes whether the individual was randomly assigned to participate ($Z = 1$) or not ($Z = 0$) in the Job Corps program. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001). Numbers in parentheses are standard errors. * and ** denote statistical significance at the 10% and 5% level, respectively.

Table 2: Estimates of Selected Point Identified Objects

	Labor Market Outcomes Sample			Public Benefits Sample	
	Enrollment	Earnings	Employment	Enrollment	Public benefits
Average for $Z = 1$	.730**	228.78**	.608**	.732**	747.21**
	(.006)	(3.004)	(.006)	(.005)	(23.40)
Average for $Z = 0$	.043**	206.60**	.570**	.043**	831.50**
	(.003)	(3.552)	(.008)	(.003)	(30.28)
ITT	.687**	22.19**	.038**	.689**	-84.29**
	(.006)	(4.652)	(.010)	(.006)	(38.27)
$LATE_c$		32.29**	.055**		-122.28**
		(7.007)	(.015)		(56.78)
Stratum Proportions (under Assumptions 1 and 4)					
π_{nt}	.270**			.268**	
	(.006)			(.006)	
π_c	.687**			.689**	
	(.007)			(.006)	
π_{at}	.043**			.043**	
	(.003)			(.003)	
Selected Point Identified Average Outcomes (under Assumptions 1 to 4)					
$E[Y(1)\|at]$		132.10**	.393**		545.45**
		(14.94)	(.037)		(110.12)
$E[Y(0)\|nt]$		223.79**	.600**		880.67**
		(5.967)	(.012)		(47.98)
$E[Y(1)\|c]$		236.82**	.624**		707.81**
		(4.022)	(.008)		(28.26)
$E[Y(0)\|c]$		204.53**	.569**		830.09**
		(5.655)	(.012)		(49.69)
$E[Y\|Z = 1, D = 1]$		230.63**	.611**		698.35**
		(3.614)	(.007)		(25.87)
$E[Y\|Z = 0, D = 0]$		209.96**	.578**		844.25**
		(3.709)	(.008)		(33.18)
Relevant Average Outcome Differences (under Assumptions 1 to 4)					
$E[Y(1)\|at] - E[Y(1)\|c]$		-104.72**	-.232**		-162.36
		(16.56)	(.040)		(119.80)
$E[Y(0)\|nt] - E[Y(0)\|c]$		19.26*	.030		50.57
		(9.902)	(.021)		(80.94)
$E[Y(1)\|at] - E[Y(0)\|nt]$		-91.70**	-.207**		-335.22**
		(16.37)	(.039)		(123.90)
$E[Y(1)\|at] - E[Y\|Z = 0, D = 0]$		-77.86**	-.185**		-298.80**
		(15.66)	(.038)		(115.62)
$E[Y\|Z = 1, D = 1] - E[Y(0)\|nt]$		6.834	.011		-182.32**
		(6.966)	(.014)		(54.59)
$E[Y(1)\|at] - E[Y(0)\|c]$		-72.43**	-.176**		-284.64**
		(16.24)	(.039)		(123.99)
$E[Y(1)\|c] - E[Y(0)\|nt]$		13.03*	.025*		-172.85**
		(7.153)	(.015)		(58.82)

Note: Z denotes whether the individual was randomly assigned to participate ($Z = 1$) or not ($Z = 0$) in the Job Corps program. D denotes whether the individual ever enrolled in the program ($D = 1$) or not ($D = 0$) during the 4 years (208 weeks) after randomization. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001). Numbers in parentheses are standard errors calculated using 5000 bootstrap repetitions. * and ** denote statistical significance at the 10% and 5% level, respectively.

Table 3: Average Baseline Characteristics of Strata in the Labor Market Outcomes Sample

Variable	nt	c	at	$nt-c$	$c-at$	$nt-at$
Female	.467**	.397**	.324**	.070**	.073**	.143**
	(.011)	(.007)	(.035)	(.015)	(.037)	(.036)
Age at Baseline	18.74**	18.32**	17.64**	.428**	.674**	1.102**
	(.052)	(.029)	(.133)	(.063)	(.137)	(.143)
White, Non-hispanic	.284**	.263**	.296**	.021*	-.033	-.012
	(.011)	(.006)	(.034)	(.013)	(.036)	(.036)
Black, Non-Hispanic	.472**	.484**	.488**	-.012	-.004	-.016
	(.012)	(.007)	(.037)	(.015)	(.039)	(.039)
Married	.035**	.016**	.005	.019**	.011**	.030**
	(.004)	(.002)	(.005)	(.005)	(.005)	(.006)
Has Child	.237**	.162**	.148**	.075**	.015	.089**
	(.010)	(.005)	(.028)	(.012)	(.030)	(.029)
Personal Education	10.27**	10.05**	9.637**	.224**	.408**	.632**
	(.035)	(.020)	(.095)	(.044)	(.101)	(.100)
Household Inc.: <3000	.267**	.255**	.187**	.012	.068**	.080**
	(.008)	(.005)	(.021)	(.010)	(.022)	(.022)
>18000	.181**	.181**	.233**	.000	-.052*	-.052*
	(.007)	(.004)	(.027)	(.009)	(.028)	(.027)
Personal Inc.: <3000	.750**	.799**	.843**	-.049**	-.044	-.093**
	(.010)	(.005)	(.026)	(.012)	(.027)	(.027)
>9000	.042**	.030**	.015*	.012*	.015*	.027**
	(.005)	(.002)	(.008)	(.006)	(.009)	(.009)
Have Job at Baseline	.224**	.208**	.216**	.015	-.008	.008
	(.010)	(.006)	(.031)	(.012)	(.033)	(.032)
Weekly Hours Worked	22.07**	21.29**	20.44**	.775	.853	1.629
	(.488)	(.272)	(1.652)	(.585)	(1.734)	(1.700)
Weekly Earnings	113.79**	102.76**	92.63**	11.03**	10.13	21.15**
	(2.989)	(2.041)	(7.986)	(3.989)	(8.328)	(8.562)
Had Job, Prev. Yr.	.667**	.640**	.651**	.027**	-.010	.016
	(.010)	(.006)	(.035)	(.013)	(.036)	(.035)
Months Employed, Prev.Yr.	3.684**	3.527**	3.120**	.157	.407	.563*
	(.102)	(.057)	(.310)	(.125)	(.324)	(.325)
Earnings, Prev.Yr.	3246.8**	2831.5**	2302.9**	415.30**	528.64**	943.94**
	(101.80)	(63.58)	(251.57)	(127.99)	(263.42)	(273.94)
Received Public Benefits	.607**	.588**	.596**	.020	-.009	.011
	(.011)	(.006)	(.037)	(.013)	(.038)	(.037)
Months Received Benefits	6.744**	6.503**	6.518**	.240	-.014	.226
	(.122)	(.073)	(.414)	(.153)	(.437)	(.424)

Note: Averages are estimated with the overidentified nonparametric GMM procedure described in the Appendix. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Missing values for each of the baseline variables were imputed with the mean of the variable. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001). Numbers in parentheses are standard errors. * and ** denote statistical significance at the 10% and 5% level, respectively.

Table 4: Average Baseline Characteristics of Strata in the Public Benefits Sample

Variable	nt	c	at	$nt - c$	$c - at$	$nt - at$
Female	.464**	.396**	.330**	.069**	.066*	.134**
	(.011)	(.006)	(.035)	(.014)	(.037)	(.036)
Age at Baseline	18.75**	18.31**	17.68**	.435**	.635**	1.070**
	(.049)	(.027)	(.126)	(.061)	(.135)	(.133)
White, Non-hispanic	.289**	.265**	.289**	.024*	-.024	-.000
	(.011)	(.006)	(.035)	(.014)	(.037)	(.036)
Black, Non-Hispanic	.461**	.480**	.503**	-.019	-.023	-.042
	(.012)	(.007)	(.037)	(.015)	(.039)	(.039)
Married	.036**	.016**	.006	.020**	.010**	.030**
	(.004)	(.002)	(.005)	(.005)	(.005)	(.006)
Has Child	.234**	.163**	.164**	.072**	-.001	.071**
	(.009)	(.005)	(.029)	(.012)	(.031)	(.030)
Personal Education	10.27**	10.05**	9.663**	.225**	.382**	.607**
	(.034)	(.020)	(.091)	(.043)	(.096)	(.094)
Household Inc.: <3000	.262**	.253**	.198**	.009	.055**	.064**
	(.008)	(.004)	(.020)	(.010)	(.022)	(.021)
>18000	.184**	.184**	.233**	.000	-.050*	-.049*
	(.007)	(.004)	(.028)	(.009)	(.029)	(.028)
Personal Inc.: <3000	.746**	.797**	.840**	-.051**	-.043*	-.094**
	(.010)	(.005)	(.024)	(.012)	(.026)	(.025)
>9000	.042**	.030**	.015**	.012**	.015*	.027**
	(.005)	(.002)	(.007)	(.006)	(.008)	(.009)
Have Job at Baseline	.227**	.211**	.213**	.016	-.002	.014
	(.010)	(.005)	(.028)	(.012)	(.030)	(.029)
Weekly Hours Worked	21.80**	21.41**	20.63**	.392	.774	1.165
	(.460)	(.291)	(1.426)	(.594)	(1.548)	(1.494)
Weekly Earnings	112.60**	103.55**	94.21**	9.025**	9.342	18.37**
	(2.890)	(2.180)	(7.394)	(4.094)	(7.954)	(7.804)
Had Job, Prev. Yr.	.667**	.642**	.668**	.025*	-.026	-.001
	(.011)	(.006)	(.031)	(.013)	(.033)	(.032)
Months Employed, Prev.Yr.	3.644**	3.553**	3.060**	.091	.492	.584*
	(.103)	(.057)	(.282)	(.130)	(.302)	(.299)
Earnings, Prev.Yr.	3241.9**	2863.6**	2390.4**	378.31**	473.14*	851.45**
	(99.19)	(65.20)	(233.19)	(130.21)	(250.73)	(249.72)
Received Public Benefits	.601**	.581**	.583**	.020	-.001	.019
	(.010)	(.006)	(.033)	(.013)	(.035)	(.034)
Months Received Benefits	6.684**	6.433**	6.395**	.251	.038	.289
	(.122)	(.076)	(.378)	(.158)	(.408)	(.385)

Note: Averages are estimated with the overidentified nonparametric GMM procedure described in the Appendix. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Missing values for each of the baseline variables were imputed with the mean of the variable. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001). Numbers in parentheses are standard errors. * and ** denote statistical significance at the 10% and 5% level, respectively.

Table 5: Estimated Bounds on the Population Average Treatment Effects (ATE)

	Weekly Earnings		Employment		Public Benefits	
	LB	*UB*	*LB*	*UB*	*LB*	*UB*
Bounds under Assumptions 1 to 4 and Bounded Outcome Assumption (A5)						
Proposition 1	-69.86	201.02	-.150	.163	-632.86	1812.4
	[-78.34, 210.61]		[-.167, .179]		[-702.21, 1901.6]	
Bounds Adding Monotonicity of Local Average Outcomes Assumption (A6)						
Proposition 2	22.19	201.02	.038	.163	-632.86	-84.29
	[14.18, 210.61]		[.021, .179]		[-702.21, -22.13]	
Bounds Adding Different Mean Dominance Assumptions (A7a, A7b, A7c)						
Proposition 3a (A7a)	-6.507	201.02	-.027	.163	-188.43	1812.4
	[-16.65, 210.61]		[-.050, .179]		[-265.90, 1901.6]	
Proposition 3b (A7b)	20.67	201.02	.033	.163	-145.90	1812.4
	[11.97, 210.61]		[.015, .179]		[-212.69, 1901.6]	
Proposition 3c (A7c)	22.57	201.02	.037	.163	-142.76	1812.4
	[13.72, 210.61]		[.019, .179]		[-210.62, 1901.6]	
Bounds Adding Both A6 and Each of A7a, A7b, and A7c						
Proposition 4 (A6, A7a)	20.43	201.02	.034	.163	-188.43	-84.29
	[13.01, 210.58]		[.018, .180]		[-265.95, -22.09]	
Proposition 5 (A6, A7b)	22.97	201.01	.039	.163	-145.90	-84.29
	[14.53, 210.56]		[.020, .180]		[-213.01, -21.83]	
Proposition 6 (A6, A7c)	24.61	201.04	.042	.163	-142.76	-84.29
	[16.01, 210.59]		[.023, .180]		[-210.62, -22.13]	

Note: Outcomes are measured four years after randomization. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. The bounds that do not involve minimum or maximum operators are estimated with sample analog estimators, and the confidence intervals (in square brackets) for the true value of the parameter are obtained with the Imbens and Manski (2004) procedure. For the bounds that involve minimum or maximum operators, the table shows half-median unbiased estimates of the bounds and 95% confidence intervals (in square brackets) for the true value of the parameter, both based on the method proposed by Chernozhukov, Lee, and Rosen (2013). This method is implemented using 5000 bootstrap replications for the variance-covariance matrix of the estimated bounding functions, and $100,000$ draws from a normal distribution. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001).

Table 6: Estimated Bounds on the Average Treatment Effect on the Treated (ATT)

	Weekly Earnings		Employment		Public Benefits	
	LB	UB	LB	UB	LB	UB
Bounds under Assumptions 1 to 4 and Bounded Outcome Assumption (A5)						
Proposition 1'	-53.53	43.47	-.019	.093	-921.15	-48.23
	[-67.48, 54.03]		[-.042, .116]		[-1045.36, 35.96]	
Bounds Adding Monotonicity of Local Average Outcomes Assumption (A6)						
Proposition 2'	28.67	43.47	.049	.093	-921.15	-108.72
	[18.32, 54.03]		[.027, .116]		[-1045.36, -25.63]	
Bounds Adding Different Mean Dominance Assumptions (A7a, A7b, A7c)						
Proposition 3'a (A7a)	18.39	43.47	.026	.093	-145.90	-48.23
	[8.32, 54.03]		[.005, .116]		[-227.18, 36.06]	
Proposition 3'b (A7b)	19.94	43.47	.028	.093	-141.86	-48.23
	[9.03, 54.03]		[.006, .116]		[-230.62, 36.16]	
Proposition 3'c (A7c)	20.55	43.47	.029	.093	-140.29	-48.23
	[9.28, 54.03]		[.006, .116]		[-232.15, 36.21]	
Bounds Adding Both A6 and Each of A7a, A7b, and A7c						
Proposition 4' (A6, A7a)	27.86	43.48	.047	.093	-145.90	-108.72
	[16.00, 55.65]		[.023, .119]		[-230.68, -21.95]	
Proposition 5' (A6, A7b)	27.96	43.48	.047	.093	-141.86	-108.72
	[16.08, 55.67]		[.023, .119]		[-235.62, -20.73]	
Proposition 6' (A6, A7c)	27.98	43.47	.047	.093	-140.29	-108.72
	[16.11, 55.64]		[.023, .119]		[-237.89, -20.18]	

Note: Outcomes are measured four years after randomization. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. The bounds that do not involve minimum or maximum operators are estimated with sample analog estimators, and the confidence intervals (in square brackets) for the true value of the parameter are obtained with the Imbens and Manski (2004) procedure. For the bounds that involve minimum or maximum operators, the table shows half-median unbiased estimates of the bounds and 95% confidence intervals (in square brackets) for the true value of the parameter, both based on the method proposed by Chernozhukov, Lee, and Rosen (2013). This method is implemented using 5000 bootstrap replications for the variance-covariance matrix of the estimated bounding functions, and $100,000$ draws from a normal distribution. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001).

Table 7: Estimated Bounds on $LATE$s of Never Takers and Always Takers

	Weekly Earnings				Employment				Public Benefits			
	$LATE_{nt}$		$LATE_{at}$		$LATE_{nt}$		$LATE_{at}$		$LATE_{nt}$		$LATE_{at}$	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
Bounds under Assumptions 1 to 4 and Bounded Outcome Assumption (A5)												
Proposition 1"	-223.79	641.87	-733.57	132.10	-.600	.400	-.607	.393	-880.67	6990.6	-7325.8	545.45
	[-233.61, 651.70]		[-758.51, 157.03]		[-.620, .421]		[-.668, .454]		[-959.40, 7069.3]		[-7506.9, 726.53]	
Bounds Adding Monotonicity of Local Average Outcomes Assumption (A6)												
Proposition 2"	.000	641.87	.000	132.10	.000	.400	.000	.393	-880.67	.000	-7325.8	.000
	[.000, 651.70]		[.000, 157.03]		[.000, .421]		[.000, .454]		[-959.40, .000]		[-7506.9, .000]	
Bounds Adding Different Mean Dominance Assumptions (A7a, A7b, A7c)												
Proposition 3"a (A7a)	-91.70	641.87	-91.70	132.10	-.207	.400	-.207	.393	-335.22	6990.6	-335.22	545.45
	[-118.63, 651.70]		[-118.63, 157.03]		[-.272, .421]		[-.272, .454]		[-531.16, 7069.3]		[-531.16, 726.53]	
Proposition 3"b (A7b)	6.834	641.87	-77.86	132.10	.011	.400	-.185	.393	-182.32	6990.6	-298.80	545.45
	[-4.625, 651.70]		[-103.62, 157.03]		[-.013, .421]		[-.247, .454]		[-272.08, 7069.3]		[-488.95, 726.53]	
Proposition 3"c (A7c)	13.03	641.87	-72.43	132.10	.025	.400	-.176	.393	-172.85	6990.6	-284.64	545.45
	[1.260, 651.70]		[-99.15, 157.03]		[.000, .421]		[-.240, .454]		[-264.71, 7069.3]		[-484.84, 726.53]	
Bounds Adding Both A6 and Each of A7a, A7b, and A7c												
Proposition 4" (A6, A7a)	.000	641.88	.000	132.08	.000	.400	.000	.393	-335.22	.000	-335.22	.000
	[.000, 651.70]		[.000, 158.98]		[.000, .421]		[.000, .458]		[-531.16, .000]		[-531.16, .000]	
Proposition 5" (A6, A7b)	3.032	641.89	.000	132.12	.003	.400	.000	.393	-182.32	.000	-298.80	.000
	[.000, 651.70]		[.000, 159.07]		[.000, .421]		[.000, .458]		[-272.08, .000]		[-488.95, .000]	
Proposition 6" (A6, A7c)	9.119	641.88	.000	132.12	.017	.401	.000	.393	-172.85	.000	-284.64	.000
	[.000, 651.70]		[.000, 159.07]		[.000, .421]		[.000, .458]		[-264.71, .000]		[-484.96, .000]	

Note: Outcomes are measured four years after randomization. Public benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. The bounds that do not involve minimum or maximum operators are estimated with sample analog estimators, and the confidence intervals (in square brackets) for the true value of the parameter are obtained with the Imbens and Manski (2004) procedure. For the bounds that involve minimum or maximum operators, the table shows half-median unbiased estimates of the bounds and 95% confidence intervals (in square brackets) for the true value of the parameter, both based on the method proposed by Chernozhukov, Lee, and Rosen (2013). This method is implemented using 5000 bootstrap replications for the variance-covariance matrix of the estimated bounding functions, and $100,000$ draws from a normal distribution. Computations use weights that account for sample design, interview design, and interview non-response (Schochet, 2001). For weekly earnings, the 90% confidence interval on $LATE_{nt}$ under Proposition 6" equals [1.362, 649.53]. For employment, the 90% confidence intervals on $LATE_{nt}$ under Propositions 3"b and 6" equal [0.006, 0.416] and [0.001, 0.416], respectively.

References

[1] Angrist, J. and Fernandez-Val, I. (2010), "Extrapolate-ing: external validity and overidentification in the $LATE$ framework," Working Paper 16566, NBER.

[2] Angrist, J., Imbens, G., and Rubin, D. (1996), "Identification of causal effects using instrumental variables," Journal of the American Statistical Association 91, 444-472.

[3] Balke, A. and Pearl, J. (1997), "Bounds on treatment effects from studies with imperfect compliance," Journal of the American Statistical Association 92(439), 1171-1176.

[4] Bampasidou, M., Flores, C., Flores-Lagunes, A., and Parisian, D. (2014), "The role of degree attainment in the differential impact of Job Corps on adolescents and young adults," Research in Labor Economics 40: 113-156.

[5] Bhattacharya, J., Shaikh, A., and Vytlacil, E. (2008), "Treatment effect bounds under monotonicity assumptions: an application to Swan-Ganz catheterization," American Economic Review: Papers & Proceedings 98:2, 351-356.

[6] Blanco, G., Flores, C. and Flores-Lagunes, A. (2012), "Bounds on average and quantile treatment effects of Job Corps training on wages," Journal of Human Resources 48 (3): 659-701.

[7] Burghardt, J., Schochet, P.Z., McConnell, S., Johnson, T., Gritz, R.M., Glazerman, S., Homrighausen, J. and Jackson, R. (2001), "Does Job Corps work? Summary of the National Job Corps Study," 8140-530, Mathematica Policy Research, Inc., Princeton, NJ.

[8] Chen, X. and Flores, C. (2014),"Bounds on Treatment Effects in the Presence of Sample Selection and Noncompliance: The Wage Effects of Job Corps", Journal of Business and Economic Statistics, forthcoming.

[9] Chernozhukov, V., Lee, S. and Rosen, A. (2013), "Intersection bounds: estimation and inference," Econometrica 81 (2): 667-737.

[10] Chesher, A. (2010), "Instrumental variable models for discrete outcomes," Econometrica 78 (2): 575-601.

[11] Chiburis, R. (2010a), "Bounds on treatment effects using many types of monotonicity," Working paper.

[12] Chiburis, R. (2010b), "Semiparametric bounds on treatment effects," Journal of Econometrics 159, 267-275.

[13] Deaton, A. (2010), "Instruments, randomization, and learning about development," Journal of Economic Literature 48, 424-455.

[14] Dehejia, R. and Wahba, S. (1999), "Causal effects in nonexperimental studies: reevaluating the evaluation of training programs," Journal of the American Statistical Association 94, 1053-1062.

[15] Flores, C. and Flores-Lagunes, A. (2010), "Nonparametric partial identification of causal net and mechanism average treatment effects ," Working paper. California Polytechnic State University, San Luis Obispo, CA.

[16] Flores, C. and Flores-Lagunes, A. (2013), "Partial identification of local average treatment effects with an invalid instrument," Journal of Business and Economic Statistics 31 (4): 534-545.

[17] Flores-Lagunes, A., Gonzalez, A. and Neumann, T. (2010), "Learning but not earning? The impact of Job Corps training on Hispanic youth," Economic Inquiry 48 (3): 651-667.

[18] Frumento, F., Mealli, F., Pacini, B. and Rubin, D. (2012), "Evaluating the Effect of Training on Wages in the Presence of Noncompliance, Nonemployment, and Missing Outcome Data," Journal of the American Statistical Association, 107 (498), 450-466.

[19] Hahn, J. (2010), "Bounds on ATE with discrete outcomes," Economics Letters 109, 24-27.

[20] Heckman, J. (1996), "On air: Identification of causal effects using instrumental variables," Journal of The American Statistical Association, 91, 459-462 .

[21] Heckman, J. (2010), "Building bridges between structural and program evaluation approaches to evaluating policy," Journal of Economic Literature, 48 (2), 356-398.

[22] Heckman, J., LaLonde, R. and Smith, J. (1999), "The economics and econometrics of active labor market programs," in Handbook of Labor Economics 3A, 1865-2097, ed. by Ashenfelter, O. and Card, D., North-Holland.

[23] Heckman, J. and Urzua, S. (2010), "Comparing IV with structural models: What simple IV can and cannot Identify," Journal of Econometrics, 156 (1), 27–37.

[24] Heckman, J. and Vytlacil, E. (2000), "Instrumental variables, selection models, and tight bounds on the average treatment effect," Technical Working Paper 259, NBER.

[25] Hirano, K. and Porter, J. (2012), "Impossibility results for nondifferentiable functionals," Econometrica 80 (4): 1769-1790.

[26] Huber, M., Laffers, L. and Mellace, G. (2015), "Sharp IV bounds on average treatment effects on the treated and other populations under endogeneity and noncompliance," Journal of Applied Econometrics, forthcoming.

[27] Huber, M. and Mellace, G. (2010), "Sharp IV bounds on average treatment effects under endogeneity and noncompliance," Discussion Paper no. 2010-31, Universität St. Gallen.

[28] Imbens, G. and Angrist, J. (1994), "Identification and estimation of local average treatment effects," Econometrica 62 (2), 467-475.

[29] Imbens, G. and Manski, C. (2004), "Confidence intervals for partially identified parameters", Econometrica 72 (6), 1845-1857.

[30] Kitagawa, T. (2009), "Identification region of the potential outcome distributions under instrument independence," CEMMAP working paper.

[31] LaLonde, R. (1986), "Evaluating the econometric evaluations of training programs with experimental data," American Economic Review 76(4), 604-620.

[32] Lee, D. (2009), "Training, wages, and sample selection: estimating sharp bounds on treatment effects," Review of Economic Studies 76, 1071-1102.

[33] Machado, C., Shaikh, A. and Vytlacil, E. (2009), "Instrumental variables and the sign of the average treatment effect," Working paper.

[34] Manski, C. (1990) "Nonparametric bounds on treatment effects," American Economic Review: Papers and Proceedings 80, 319-323.

[35] Manski, C. (1997), "Monotone treatment response," Econometrica 65(6), 1311-1334.

[36] Manski, C. and Pepper, J. (2000), "Monotone instrumental variables: with an application to the returns to schooling," Econometrica 68 (4), 997-1010.

[37] Moffitt, R. (2003), "The Temporary Assistance for Needy Families Program," in Means-Tested Transfer Programs in the United States, 291-363, ed. by Moffitt, R., University of Chicago Press.

[38] Robins, J. and Greenland, S. (1996), "Comment on Angrist, Imbens and Rubin: Estimation of the global average treatment effects using instrumental variables," Journal of the American Statistical Association, 91, 456-458.

[39] Schochet, P., Burghardt, J. and Glazerman, S., 2001. National Job Corps Study: the impacts of Job Corps on participants' employment and related outcomes. Mathematica Policy Research, Inc., Princeton, NJ.

[40] Schochet, P., Burghardt, J. and McConnell, S. (2008), "Does Job Corps work? Impact findings from the National Job Corps Study," American Economic Review, 98(5), 1864-1886.

[41] Shaikh, A. and Vytlacil, E. (2011), "Partial identification in triangular systems of equations with binary dependent variables," Econometrica 79(3), 949-955.

[42] U.S. Department of Labor (2015), "Job Corps fact sheet," http://www.doleta.gov/Programs/factsht/jobcorps.cfm (Accessed March 10, 2015).

[43] van Ours, J. (2004), "The Locking-in Effect of Subsidized Jobs," Journal of Comparative Economics, 32, 37-55.

[44] Vytlacil, E. (2002), "Independence, monotonicity, and latent index models: an equivalence result," Econometrica 70 (1), 331-341.

[45] Zhang, J. Rubin, D., and Mealli, F. (2008), "Evaluating the effects of job training programs on wages through principal stratification," In D. Millimet et al. (eds) Advances in Econometrics vol XXI, Elsevier.

A Appendix

A.1 Proof

We present only the proof of Proposition 2, as the proofs for the rest of the propositions are similar. Under Assumptions 1 through 4, AIR show that $LATE_c = (E[Y|Z=1] - E[Y|Z=0])/(p_{1|1} - p_{1|0})$. By Assumption 6(ii), and since we have ordered Z such that $p_{1|1} > p_{1|0}$, the direction of the monotonicity in Assumption 6(i) is identified from the sign of $LATE_c$. Here we consider only the case when $LATE_c > 0$, as the sharp bounds when $LATE_c < 0$ are constructed in the same way. From equation (1) we can write $ATE = \pi_{at}(E[Y(1)|at] - E[Y(0)|at]) + \pi_{nt}(E[Y(1)|nt] - E[Y(0)|nt]) + \pi_c LATE_c$. Under Assumptions 1 though 4, the sampling process identifies each of the quantities to the right of this equation except for $E[Y(1)|nt]$ and $E[Y(0)|at]$, and thus equation (2) follows. Since there are no restrictions on these two means other than those imposed by Assumptions 5 and 6(i), these two assumptions directly imply the bounds $y^u \geq E[Y(1)|nt] \geq E[Y(0)|nt] = \overline{Y}^{10}$ and $\overline{Y}^{01} = E[Y(1)|at] \geq E[Y(0)|at] \geq y^l$. The lower (upper) bound on ATE in Proposition 2 is obtained from equation (2) by setting $E[Y(1)|nt]$ at its lower (upper) bound and $E[Y(0)|at]$ at its upper (lower) bound.

For sharpness, first, ATE attains its smallest value when $E[Y(0)|at] = \overline{Y}^{01}$ and $E[Y(1)|nt] = \overline{Y}^{10}$. Otherwise, always-takers or never-takers violate Assumption 6(i). Similarly, ATE attains its largest value when $E[Y(0)|at] = y^l$ and $E[Y(1)|nt] = y^u$. Otherwise, always-takers or never-takers violate Assumption 5. Next, we will show that $\forall \alpha \in [LB, UB]$, there exist distributions consistent with observed data, and $ATE = \alpha$ evaluated under such distributions. $\forall \alpha \in [LB, UB]$, it can be written as $\alpha = \overline{Y}^{11} p_{1|1} - \overline{Y}^{00} p_{0|0} + q_1 p_{0|1} - q_0 p_{1|0}$, where $q_1 \in [\overline{Y}^{10}, y^u]$ and $q_0 \in [y^l, \overline{Y}^{01}]$. Let $F_{Y_1|Z,D}(y_1|1,d)$ denote the distribution of the potential outcome $Y(1)$ conditional on $Z=1$ and $D=d$. Similarly, $F_{Y_0|Z,D}(y_0|0,d)$ denotes the distribution of the potential outcome $Y(0)$ conditional on $Z=0$ and $D=d$. Then, define

$$F_{Y_1|Z,D}(y_1|1,d) = \begin{cases} F_{Y|Z,D}(y|1,1), & \text{if } D=1 \\ 1[y_1 \geq q_1], & \text{if } D=0 \end{cases}$$

and

$$F_{Y_0|Z,D}(y_0|0,d) = \begin{cases} F_{Y|Z,D}(y|0,0), & \text{if } D=0 \\ 1[y_0 \geq q_0], & \text{if } D=1 \end{cases}.$$

$ATE = E[Y(1) - Y(0)]$
$= E[Y(1)|Z=1] - E[Y(0)|Z=0]$
$= p_{1|1}E[Y(1)|Z=1, D=1] + p_{0|1}E[Y(1)|Z=1, D=0] - p_{1|0}E[Y(0)|Z=0, D=1] - p_{0|0}E[Y(0)|Z=0, D=0]$
$= p_{1|1}E[Y|Z=1, D=1] + p_{0|1}E[Y(1)|Z=1, D=0] - p_{1|0}E[Y(0)|Z=0, D=1] - p_{0|0}E[Y|Z=0, D=0]$
$= p_{1|1}\overline{Y}^{11} + p_{0|1}q_1 - p_{1|0}q_0 - p_{0|0}\overline{Y}^{00}$
$= \alpha.$

The second line follows Assumption 1, the third line follows Law of Iterated Expectation, and the fourth and fifth lines follow the defined distributions.

A.2 GMM Moment Function

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

$$g(\{\overline{x}_k\}) = \begin{bmatrix} (x - \overline{x}_{at})(1-Z)D \\ (x - \overline{x}_{nt})Z(1-D) \\ (x - \overline{x}_c \frac{\pi_c}{p_{1|1}} - \overline{x}_a \frac{\pi_{at}}{p_{1|1}})ZD \\ (x - \overline{x}_c \frac{\pi_c}{p_{0|0}} - \overline{x}_n \frac{\pi_{nt}}{p_{0|0}})(1-Z)(1-D) \\ x - \sum_k \pi_k \overline{x}_k \end{bmatrix}$$

where $\{\overline{x}_k\} = \{\overline{x}_{at}, \overline{x}_{nt}, \overline{x}_c\}$. By Law of Iterated Expectation, $E[g(\{\overline{x}_k\})] = 0$ when evaluated at the true value of $\{\overline{x}_k\}$.

Alternatively, we could also write the moment function for the proportions of all the strata and then estimate the model together with the average baseline characteristics simultaneously by GMM. However, such GMM estimators do not behave well in our data. Thus, in our application, we first identify the proportions of all the strata, and then estimate all the average baseline characteristics given the identified proportions. As seen in $g(\{\overline{x}_k\})$, for each variable, we have 5 equations (4 derived from the conditional expectations defined by $\{Z, D\}$ plus one from the expectation for the entire sample) to identify 3 means, i.e., $\{\overline{x}_k\}$. Since the standard errors obtained from this GMM model do not take into account the fact that the proportions of the strata are also estimated, we employ a 500-repetition bootstrap to calculate the standard errors of the estimated average baseline characteristics.

A.3 Bounds on $LATE_{nt}$ and $LATE_{at}$

This subsection presents the bounds on $LATE_{nt}$ and $LATE_{at}$ under each set of assumptions considered in the paper. They are obtained by using the equations: $LATE_{nt} = E[Y(1)|nt] - \overline{Y}^{10}$ and $LATE_{at} = \overline{Y}^{01} - E[Y(0)|at]$. The propositions below, labeled Proposition 1" to Proposition 6", show the bounds for those two parameters and are analogous to those presented in the main text for the ATE and ATT.

Proposition 1" *Under Assumptions 1 through 5, sharp bounds on* $LATE_{nt}$ *and* $LATE_{at}$ *are given by:* $y^l - \overline{Y}^{10} \leq LATE_{nt} \leq y^u - \overline{Y}^{10}$, *and* $\overline{Y}^{01} - y^u \leq LATE_{at} \leq \overline{Y}^{01} - y^l$.

Proposition 2" *Under Assumptions 1 through 6, sharp bounds on* $LATE_{nt}$ *and* $LATE_{at}$ *are given by: If* $E[Y|Z=1] - E[Y|Z=0] > 0$, $0 \leq LATE_{nt} \leq y^u - \overline{Y}^{10}$, *and* $0 \leq LATE_{at} \leq \overline{Y}^{01} - y^l$; *if* $E[Y|Z=1] - E[Y|Z=0] < 0$, $y^l - \overline{Y}^{10} \leq LATE_{nt} \leq 0$, *and* $\overline{Y}^{01} - y^u \leq LATE_{at} \leq 0$.

Proposition 3" *Sharp bounds on $LATE_{nt}$ and $LATE_{at}$ are given by: (a) Under Assumptions 1 through 5 and 7a,* $\overline{Y}^{01} - \overline{Y}^{10} \le LATE_{nt} \le y^u - \overline{Y}^{10}$, *and* $\overline{Y}^{01} - \overline{Y}^{10} \le LATE_{at} \le \overline{Y}^{01} - y^l$*; (b) Under Assumptions 1 through 5 and 7b,* $\overline{Y}^{11} - \overline{Y}^{10} \le LATE_{nt} \le y^u - \overline{Y}^{10}$, *and* $\overline{Y}^{01} - \overline{Y}^{00} \le LATE_{at} \le \overline{Y}^{01} - y^l$*; (c) Under Assumptions 1 through 5 and 7c,*

$$\begin{aligned} \frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}} - \overline{Y}^{10} &\le LATE_{nt} \le y^u - \overline{Y}^{10}, \\ \overline{Y}^{01} - \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}} &\le LATE_{at} \le \overline{Y}^{01} - y^l. \end{aligned}$$

Proposition 4" *Under Assumptions 1 through 6 and 7a, sharp bounds on $LATE_{nt}$ and $LATE_{at}$ are given by: If* $E[Y|Z=1] - E[Y|Z=0] > 0$, $\max\{\overline{Y}^{01}, \overline{Y}^{10}\} - \overline{Y}^{10} \le LATE_{nt} \le y^u - \overline{Y}^{10}$, *and* $\overline{Y}^{01} - \min\{\overline{Y}^{01}, \overline{Y}^{10}\} \le LATE_{at} \le \overline{Y}^{01} - y^l$*; if* $E[Y|Z=1] - E[Y|Z=0] < 0$, $\overline{Y}^{01} - \overline{Y}^{10} \le LATE_{nt} \le 0$, *and* $\overline{Y}^{01} - \overline{Y}^{10} \le LATE_{at} \le 0$.

Proposition 5" *Under Assumptions 1 through 6 and 7b, sharp bounds on $LATE_{nt}$ and $LATE_{at}$ are given by: If* $E[Y|Z=1] - E[Y|Z=0] > 0$, $\max\{\overline{Y}^{11}, \overline{Y}^{10}\} - \overline{Y}^{10} \le LATE_{nt} \le y^u - \overline{Y}^{10}$, *and* $\overline{Y}^{01} - \min\{\overline{Y}^{01}, \overline{Y}^{00}\} \le LATE_{at} \le \overline{Y}^{01} - y^l$*; if* $E[Y|Z=1] - E[Y|Z=0] < 0$, $\overline{Y}^{11} - \overline{Y}^{10} \le LATE_{nt} \le 0$, *and* $\overline{Y}^{01} - \overline{Y}^{00} \le LATE_{at} \le 0$.

Proposition 6" *Under Assumptions 1 through 6 and 7c, sharp bounds on $LATE_{nt}$ and $LATE_{at}$ are given by: If* $E[Y|Z=1] - E[Y|Z=0] > 0$,

$$\begin{aligned} \max\left\{\frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}}, \overline{Y}^{10}\right\} - \overline{Y}^{10} &\le LATE_{nt} \le y^u - \overline{Y}^{10}, \\ \overline{Y}^{01} - \min\left\{\overline{Y}^{01}, \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}}\right\} &\le LATE_{at} \le \overline{Y}^{01} - y^l; \end{aligned}$$

if $E[Y|Z=1] - E[Y|Z=0] < 0$,

$$\begin{aligned} \frac{\overline{Y}^{11} p_{1|1} - \overline{Y}^{01} p_{1|0}}{p_{1|1} - p_{1|0}} - \overline{Y}^{10} &\le LATE_{nt} \le 0, \\ \overline{Y}^{01} - \frac{\overline{Y}^{00} p_{0|0} - \overline{Y}^{10} p_{0|1}}{p_{1|1} - p_{1|0}} &\le LATE_{at} \le 0. \end{aligned}$$