Partial Identification of Local Average Treatment Effects with an Invalid Instrument*

Carlos A. Flores† Alfonso Flores-Lagunes‡

March 9, 2011

Abstract

We derive nonparametric bounds for local average treatment effects without requiring the exclusion restriction assumption to hold or an outcome with a bounded support. Instead, we employ assumptions requiring weak monotonicity of mean potential outcomes within or across subpopulations defined by the values of the potential treatment status under each value of the instrument. These results are employed to bound the effect of attaining a GED, high school, or vocational degree on future employment and weekly earnings, using randomization into a training program as an invalid instrument. The resulting bounds are informative, indicating that the local effect when assigned to training for those whose degree attainment is affected by the instrument is at most 10 percentage points on employment and $54 on weekly earnings.

Key words and phrases: causal inference, instrumental variables, treatment effects, nonparametric bounds, principal stratification

JEL classification: C13, C21, C14

*Useful comments have been provided by Joshua Angrist, Chris Parmeter, and participants at the 2010 NY Camp Econometrics. This work has been partially supported by NSF under grants SES-0852211 and SES-0852139. Competent research assistance was provided by Maria Bampasidou. All errors are our own.

†Department of Economics, University of Miami. Email: caflores@miami.edu

‡Food and Resource Economics Department and Economics Department, University of Florida, and IZA, Bonn, Germany. Email: alfonsofl@ufl.edu

1 Introduction

Instrumental variable (IV) methods are widely used in economics and other fields to analyze the effect of a treatment on an outcome. These methods exploit exogenous variation in the treatment coming from exogenous variation in another variable, called the instrument. A widely used framework for studying IV methods was developed in Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996) (hereafter IA and AIR, respectively). They show that in the presence of heterogeneous effects and under some assumptions, IV estimators point identify the local average treatment effect ($LATE$), defined as the average treatment effect for those individuals whose treatment status is changed because of the instrument (known as compliers). A critical assumption of IV methods is the exclusion restriction, which in the $LATE$ framework requires that the instrument affects the outcome only through its effect on the treatment. Since this assumption is not directly testable, it is debatable in many applications whether the instrument employed is valid (i.e., satisfies the exclusion restriction). Moreover, in many relevant empirical applications it is difficult to find valid instruments. In this paper, we derive analytic nonparametric bounds for $LATE$ without imposing the exclusion restriction assumption or requiring an outcome with bounded support. Instead, we employ assumptions requiring weak monotonicity of mean potential outcomes within or across subpopulations defined by the values of the potential treatment status under each value of the instrument. In practice, the assumptions we consider can be substantiated with economic theory, combined with each other depending on their plausibility, and some of them can be falsified from the data employing their testable implications.

There is a growing literature on partial identification of treatment effects in IV models. A strand of this literature constructs nonparametric bounds on average treatment effects assuming the availability of a valid instrument (Manski, 1990, 1994; Balke and Pearl, 1997; Heckman and Vytlacil, 1999, 2000; Shaikh and Vytlacil, 2005; Bhattacharya et al., 2008). This paper is different in that our partial identification results do not require a valid instrument. Another strand of the literature considers invalid instruments and develops bounds on average treatment effects. Conley et al. (2008) use information on a parameter summarizing the extent of violation of the exclusion restriction along with distributional assumptions in the form of deterministic or probabilistic priors. Nevo and Rosen (2008) derive analytic bounds on average treatment effects by employing assumptions on the sign and extent of correlation between the instrument and the error term in a linear model. In contrast to these two papers, our approach is nonparametric in nature and does not require modeling the extent of invalidity of the instrument nor its correlation with an error term, while it is limited to the case of a binary and randomly assigned instrument and a binary endogenous regressor.

Our paper is closer in spirit to the influential work of Manski and Pepper (2000) in the sense that we also study nonparametric partial identification of average treatment effects without assuming the validity of an instrument. Manski and Pepper (2000) introduced the monotone instrumental variable (MIV) assumption and analyzed its identifying power. This assumption relaxes the traditional IV assumption, which requires equality of mean responses for subpopulations with different values of the IV, by replacing the equality with a weak inequality. Our paper differs from their work in several ways. First, we focus on deriving bounds for $LATE$, while they focus (as does most of the literature on partial identification in IV models) on the population average treatment effect. Second, our bounds do not require the outcome to have a bounded support while, in general, the bounds in Manski and Pepper (2000) are uninformative without a bounded outcome.[1] Third, in contrast to Manski and Pepper (2000) and the rest of the literature on partial identification in IV models, our bounds are derived within the principal stratification framework (Frangakis and Rubin, 2002). Principal stratification provides a framework for analyzing causal effects when controlling for a variable that has been affected by the treatment. This framework permits us to analyze causal effects when allowing the IV to causally affect the outcome through channels other than the treatment, hence allowing the exclusion restriction to be violated. Finally, as we later discuss, the assumptions considered in this paper are related to but different from those in Manski and Pepper (2000).

Deriving bounds for $LATE$ is important for several reasons. First, $LATE$ is a widely used parameter in applied work and its bounds can be employed as a robustness check when estimating it, as illustrated in our empirical application. Second, the bounds for $LATE$ can be more informative than those of other parameters (e.g., the population average treatment effect) in some settings. Finally, in some applications the average treatment effect for compliers is a relevant parameter even if the instrument is invalid, such as in experiments (or quasi-experiments) with imperfect compliance. For instance, consider the classical example in AIR, where they study the effect of military service on civilian earnings using the Vietnam era draft lottery as an IV. In this case, $LATE$ is the average effect of military service for those individuals that enrolled in the military because their draft lottery number made them eligible to be drafted. Thus, even if the draft lottery had a separate effect on civilian earnings through channels other than the military service (e.g., by affecting schooling decisions), $LATE$ is still a policy-relevant parameter as it measures the effect on those induced to the military by the draft.[2]

The setup considered, which follows the original setup in AIR, consists of a randomly as-

[1] Manski and Pepper (2000) show that in the special case when the IV is the realized treatment, informative bounds can be constructed even if the outcome has unbounded support under the MIV assumption (in this special case called "monotone treatment selection") and the monotone treatment response assumption of Manski (1997).

[2] There is considerable debate in the literature regarding the usefulness of the $LATE$ parameter (e.g., Deaton 2010a,b; Heckman and Urzua; 2010; Angrist and Pischke, 2010; Imbens, 2010). We do not argue that $LATE$ is a relevant parameter in every possible setting. If interest lies on bounding the population average treatment effect based on an invalid instrument, we refer the reader to the work by Manski and Pepper (2000, 2009).

signed binary instrument and a binary treatment. This case sets the basis for extensions to other settings and it allows us to focus on the main ideas behind our partial identification results. This setting is also important in practice. Most of the program evaluation literature focuses on the binary treatment case (e.g., Imbens and Wooldridge, 2009), and binary instruments are common in empirical applications (e.g., Rosenzweig and Wolpin, 1980; Angrist, 1990; Oreopoulous, 2006). Moreover, randomized experiments have gained importance in many fields in economics as a way of estimating average causal effects, such as in labor (e.g., Heckman et al., 1999) and development economics (e.g., Duflo et al., 2008). By applying the methods in this paper, the randomized variable in those experiments can be employed as an instrumental variable to derive bounds for relevant treatment effects even if it does not satisfy the exclusion restriction. Indeed, this is a situation we exploit in our analysis of the returns to attaining a degree. Similarly, the tools developed herein are also useful in the design of experiments in instances where it is difficult to randomize a treatment of interest. In such cases, one could randomize a variable that affects the treatment instead, and employ the methods herein to bound the effect of interest.

This work exploits the close relationship between IV models and the growing literature on mechanism and net, or indirect and direct, effects (e.g., Robins and Greenland, 1992; Pearl, 2001; Robins, 2003; Imai et al., 2010; Flores and Flores-Lagunes, 2010). Mechanism and net average treatment effects ($MATE$ and $NATE$, respectively) provide an intuitive decomposition of the population total average effect of a variable on an outcome that enables learning about the part of this effect that works through a given mechanism. We use these effects to explicitly allow the instrument to have a causal effect on the outcome net of its effect through the treatment variable. Hence, in lieu of assuming that all the effect of the instrument on the outcome works through the treatment (i.e., the exclusion restriction assumption), we let the instrument have a mechanism and a net average effect on the outcome, where the mechanism is the treatment of interest. One intuitive way to think of the part of the effect of the instrument on the outcome that works through the treatment ($MATE$) is as the reduced-form effect of the instrument on the outcome when the IV is assumed to satisfy the exclusion restriction.

More specifically, to derive bounds for $LATE$, we show that it can be written as a function of $MATE$ and the average effect of the instrument on the treatment. The latter effect is point identified, while $MATE$ can be partially identified under relatively weak assumptions (Flores and Flores-Lagunes, 2010). The approach followed to derive bounds for $MATE$ consists of writing it as a function of the mean potential outcomes in each of the subpopulations or strata where all individuals have the same values of the potential treatment status under each value of the instrument, and then imposing assumptions relating the (partially or point) identified mean potential outcomes of the different strata in the population to those that are unidentified. In particular, we consider two sets of assumptions that impose weak-inequality restrictions on the mean potential outcomes of the different strata. These sets of assumptions may be used together or separately. The first set imposes weak monotonicity of different mean potential outcomes within a given strata, while the second set imposes weak monotonicity for the same mean potential outcome across strata.

In the first part of Section 2, we define our $LATE$ and relate it to $MATE$. The $LATE$ we analyze differs from the one in IA and AIR in that in our setting, it is necessary to specify whether or not the effect of the treatment on the outcome is under exposure to the instrument. The reason for this is that we allow the instrument to have an effect on the way the treatment affects the outcome, so that the treatment effect for the same individual can be different depending on whether or not she was exposed to the instrument. In the second part of Section 2, we present the main partial identification results. Section 3 employs these methods to analyze the effect of obtaining a general educational development (GED), high school, or vocational degree on labor market outcomes using randomization into a training program (Job Corps) as an instrument. This application relates to the large empirical literature on the effect of education and degrees (i.e., credentials) on labor market outcomes (e.g., Card, 1999; Hungerford and Solon, 1987; Cameron and Heckman, 1993; Jaeger and Page, 1996; Flores-Lagunes and Light, 2010). In this application, assignment into training is not likely to satisfy the exclusion restriction assumption to point identify the effect of interest because it may affect the outcomes through channels other than the attainment of a degree (e.g., through the use of the other components of the training program, such as job search services or social skill training). Hence, we construct bounds on the $LATE$ of obtaining such a degree on employment and weekly earnings regarding assignment into training as an invalid instrument. Our results suggest that the average effect of attaining such a degree on employment (weekly earnings) when assigned to training for those individuals whose attainment of a degree is affected by this assignment is at most 10 percentage points ($53.62).

2 Partial Identification of LATE

2.1 Setup and Relation between $LATE$ and $MATE$

Assume we have a random sample of size n from a large population. For each unit i in the sample, let $D_i \in \{0,1\}$ indicate whether the unit received the treatment of interest ($D_i = 1$) or the control treatment ($D_i = 0$). The focus is on analyzing the effect of the treatment on an outcome Y. Let $Y_i(1)$ and $Y_i(0)$ denote the two potential outcomes as a function of the treatment. They represent the outcome individual i would get if she received the treatment or not, respectively. We consider employing exogenous variation in a binary variable Z to learn about the effect of D on Y, with $Z_i \in \{0,1\}$. Let $D_i(1)$ and $D_i(0)$ denote the potential treatment status; that is, the treatment status individual i would receive depending on the value of Z_i.

Similarly, we need to include Z in the definition of the potential outcomes. Let $Y_i(z,d)$ denote the potential outcome individual i would obtain if she received a value of the instrument and the treatment of z and d, respectively. For each unit i, we observe the vector (Z_i, D_i, Y_i), where $D_i = Z_i D_i(1) + (1 - Z_i) D_i(0)$ and $Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$. To simplify notation, in the rest of the paper we write the subscript i only when deemed necessary.[3]

AIR partition the population into groups such that, within each group, all individuals have the same values of the vector $\{D_i(0), D_i(1)\}$. Frangakis and Rubin (2002) call such a partition a "basic principal stratification" and note that comparisons of potential outcomes within these strata yield causal effects because the strata an individual belongs to is not affected by the value of the instrument received. Our setting gives rise to four principal strata: $\{1,1\}$, $\{0,0\}$, $\{0,1\}$ and $\{1,0\}$. These strata are commonly referred to as always takers, never takers, compliers, and defiers, respectively.

IA and AIR impose the following assumptions:

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

Assumption 2 *(Nonzero Average Effect of Z on D).* $E[D(1) - D(0)] \neq 0$.

Assumption 3 *(Individual-Level Monotonicity of Z on D).* $D_i(1) \geq D_i(0)$ *for all* i.

Assumption 2 requires the instrument to have an effect on the treatment status, while Assumption 3 rules out the existence of defiers.

In addition, IA and AIR impose the following assumption to point identify a local average treatment effect of D on Y:

$$\textit{Exclusion Restriction Assumption (AIR):}\ Y_i(0,d) = Y_i(1,d) \text{ for all } i \text{ and } d \in \{0,1\}. \tag{1}$$

This assumption requires that any effect of the instrument on the potential outcomes is through the treatment status only. Vytlacil (2002) shows that the IV assumptions imposed in the framework of IA and AIR are equivalent to those imposed in nonparametric selection models. IA and AIR show that if the exclusion restriction holds, along with Assumptions 1, 2 and 3, we can point identify the average causal effect of D on Y for the compliers:

$$E[Y(1) - Y(0) | D(1) - D(0) = 1] = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}. \tag{2}$$

IA and AIR refer to the effect in (2) as the local average treatment effect, or $LATE$. It gives the average effect of D on Y for those individuals whose treatment status is affected by the instrument (compliers).

[3] Our notation implicitly imposes the stable unit treatment value assumption (SUTVA) in AIR. This assumption implies that the individual potential outcomes are not affected by the treatment received by other individuals.

The exclusion restriction is crucial to the point identification result in (2). Intuitively, we obtain an average effect of D on Y by dividing the reduced-form effect of Z on Y by the effect of Z on D because the exclusion restriction guarantees that all of the effect of Z on Y works through D. We allow the instrument to have a causal effect on the outcome through channels other than the treatment status. To do this, we need to learn what part of the effect of Z on Y works through D (i.e., the mechanism effect) in order to use Z to learn about the effect of D on Y. Note that the concept of mechanism effect now plays the role of the reduced-form effect of Z on Y when the exclusion restriction holds.

We introduce some additional notation. Let $Y_i^z(1)$ and $Y_i^z(0)$ denote the potential outcomes as a function of the instrument, so that they give the outcome individual i would obtain if she were or were not exposed to the instrument, respectively. Hence, the reduced-form average treatment effect of the instrument on the outcome is given by $ATE_{ZY} = E[Y^z(1) - Y^z(0)]$. Note that by definition $Y_i^z(1) = Y_i(1, D_i(1))$ and $Y_i^z(0) = Y_i(0, D_i(0))$. Also, let the potential outcome $Y_i(1, D_i(0))$ represent the outcome individual i would obtain if she were exposed to the instrument but the effect of the instrument on the treatment status were blocked by keeping her treatment status at the value she would have received had she not been exposed to the instrument. Intuitively, $Y_i(1, D_i(0))$ is the potential outcome from an alternative counterfactual experiment in which the instrument is the same as the original one but blocks the effect of Z on D by holding D_i fixed at $D_i(0)$. Following Flores and Flores-Lagunes (2010), the mechanism average treatment effect, or $MATE$, is given by

$$MATE = E[Y^z(1) - Y(1, D(0))], \tag{3}$$

and the net average treatment effect, or $NATE$, is given by

$$NATE = E[Y(1, D(0)) - Y^z(0)]. \tag{4}$$

By construction, $ATE_{ZY} = MATE + NATE$. Hence, $MATE$ and $NATE$ decompose the total average effect of the instrument on the outcome into the part that works through the treatment status ($MATE$) and the part that is net of the treatment-status channel ($NATE$). Since $Y^z(1) = Y(1, D(1))$, $MATE$ gives the average effect on the outcome from a change in the treatment status that is due to the instrument, holding the value of the instrument fixed at one. Similarly, since $Y^z(0) = Y(0, D(0))$, $NATE$ gives the average effect of the instrument on the outcome when the treatment status of every individual is held constant at $D_i(0)$. Note also that $ATE_{ZY} = MATE$ if all the effect of Z on Y works through D (i.e., if the exclusion restriction in (1) is satisfied), and $ATE_{ZY} = NATE$ if none of the effect of Z on Y works through D (either because Z does not affect D or because D does not affect Y).

Our goal is to partially identify relevant average effects of D on Y while allowing Z to have a net or direct effect on Y; thus, our first step consists of relating an average effect of D on Y

to $MATE$. Note that we can write $MATE$ in (3) as:

$$\begin{aligned} MATE &= E[Y(1, D(1)) - Y(1, D(0))] \\ &= E\{[D(1) - D(0)] \cdot [Y(1,1) - Y(1,0)]\} \\ &= \Pr(D(1) - D(0) = 1) \cdot E[Y(1,1) - Y(1,0)|D(1) - D(0) = 1] \\ &\quad - \Pr(D(1) - D(0) = -1) \cdot E[Y(1,1) - Y(1,0)|D(1) - D(0) = -1]. \end{aligned} \tag{5}$$

The second line in (5) writes $MATE$ as the expected value of the product of the individual effect of the instrument on the treatment status times the individual effect from a change in the treatment status on the outcome, holding the value of the instrument fixed at one. The third line uses iterated expectations and sets the basis for the following proposition.

Proposition 1 *Under Assumptions 2 and 3 we can write*

$$LATE \equiv E[Y(1,1) - Y(1,0)|D(1) - D(0) = 1] = \frac{MATE}{E[D(1) - D(0)]}. \tag{6}$$

Proposition 1 follows directly from (5) by ruling out the existence of defiers. As in IA and AIR, we refer to the parameter $E[Y(1,1) - Y(1,0)|D(1) - D(0) = 1]$ as the local average treatment effect ($LATE$). It gives the average treatment effect for compliers under exposure to the instrument. Proposition 1 writes $LATE$ as a function of $MATE$ and the average effect of the instrument on the treatment status. Flores and Flores-Lagunes (2010) derive nonparametric bounds for $MATE$ in a setting analogous to the one presented here. Given that the denominator in (6) is point identified under random assignment of the instrument, those bounds can be used to derive bounds for $LATE$.

To gain intuition on the result in Proposition 1, it is helpful to relate it to the corresponding results in IA (Theorem 1) and AIR (Proposition 1). Imposing the exclusion restriction assumption in (1) has two important effects. First, assuming that all the effect of the instrument on the outcome works through the treatment implies that $MATE = ATE_{ZY}$, with ATE_{ZY} being point identified under Assumption 1. Hence, the result in (6) reduces to those in IA and AIR under the exclusion restriction assumption. Second, the exclusion restriction implies that $E[Y(1,1) - Y(1,0)|D(1) - D(0) = 1] = E[Y(0,1) - Y(0,0)|D(1) - D(0) = 1] = E[Y(1) - Y(0)|D(1) - D(0) = 1]$. Intuitively, it implies that the instrument does not affect how the treatment affects the outcome. Therefore, specifying whether the effect of the treatment on the outcome is under exposure of the instrument is irrelevant. In our setting, however, this distinction is important because we allow the instrument to have a net or direct effect on the outcome, so average treatment effects can be different depending on whether or not the individuals are exposed to the instrument. As a result, the $LATE$ in (6) is not the same as that in IA and AIR without further assumptions.[4]

Just as in the IV model studied in IA and AIR, the specific instrument employed is crucial in interpreting the $LATE$ in (6). For instance, consider a case where individuals are randomly assigned to either enroll (group A) or not enroll (group B) into a training program but there is imperfect compliance; that is, some of the individuals in group A do not enroll in the program while some in group B do. If we use random assignment into groups A and B as an instrument to learn about the effect of a given treatment D on an outcome, $LATE$ in (6) would be interpreted as the average effect of D on the outcome for compliers when assigned into group A (a type of "intention to treat"), and not as the average effect of D on the outcome for compliers when enrolled into the training program.

Finally, we note that it is also possible to define $LATE$ as the average treatment effect for compliers under no exposure to the instrument, $E\left[Y(0,1) - Y(0,0)|D(1) - D(0) = 1\right]$, by using the potential outcome $Y_i(0, D(1))$ instead of $Y_i(1, D(0))$ in the definition of $MATE$ and $NATE$ above. The same approach used in this paper could be used to bound that parameter.

2.2 Bounds on LATE

In this subsection, we derive bounds on $LATE$ in (6) based on Proposition 1 and the bounds on $MATE$ derived by Flores and Flores-Lagunes (2010) within a principal stratification framework. We start by discussing partial identification of $MATE$ because, once combined with point identification of $E\left[D(1) - D(0)\right]$, partial identification of $LATE$ follows from Proposition 1.

Partial identification of $MATE$ is attained from the level of the strata up, and thus we define local versions of $MATE$ and $NATE$ as the corresponding average effects within strata. To simplify notation, we write at, nt, c and d to refer to the strata of always takers, never takers, compliers and defiers, respectively. Let

$$LMATE_k = E[Y^z(1)|k] - E[Y(1, D(0))|k], \text{ for } k = at, nt, c, d \tag{7}$$

and

$$LNATE_k = E[Y(1, D(0))|k] - E[Y^z(0)|k], \text{ for } k = at, nt, c, d. \tag{8}$$

The fact that $D_i(0) = D_i(1)$ for the always and never takers implies that for these two strata $Y^z(1) = Y(1, D(0))$, so $LMATE_k = 0$ and $LNATE_k = E[Y^z(1) - Y^z(0)|k]$ for $k = at, nt$. It also implies that the observed data contains information on $Y_i(1, D(0))$ only for those treated individuals in the nt and at strata. In addition, note that $LATE$ in (6) equals

[4] One assumption that would make $LATE$ in (6) equal to the $LATE$ in IA and AIR (see equation (2)) is that $E[Y(1,1) - Y(1,0)|D(1) - D(0) = 1] = E[Y(0,1) - Y(0,0)|D(1) - D(0) = 1]$. A stronger assumption, but still weaker than the exclusion restriction, is that $Y_i(1,1) - Y_i(1,0) = Y_i(0,1) - Y_i(0,0)$ for all i. These two assumptions allow the instrument to have an effect on the outcome but not on the effect of the treatment status on the outcome.

the local mechanism average treatment effect for compliers ($LMATE_c$), since $LMATE_c = E[Y(1, D(1)) - Y(1, D(0))|c] = E\{[D(1) - D(0)] \cdot [Y(1,1) - Y(1,0)]|c\} = E[Y(1,1) - Y(1,0)|c]$.

Under Assumptions 1 and 3, it is possible to point identify the proportion of each of the strata in the population and to point or partially identify the mean potential outcomes and local effects of certain strata. Consider the following table summarizing the relationship between the "compliance behavior" of the individuals in the sample and their observed treatment status (D_i) and instrument exposure (Z_i) under Assumption 3:

Table 1

		Z_i = 0	Z_i = 1
D_i	0	nt, c	nt
	1	at	at, c

Let π_{nt}, π_{at}, π_c, and π_d be the population proportions of each of the principal strata nt, at, c and d, respectively, and let $p_{d|z} \equiv \Pr(D_i = d|Z_i = z)$ for $d, z = 0, 1$. Then, Assumptions 1 and 3 imply that the proportions of each of the strata in the population are point identified as $\pi_{nt} = p_{0|1}$, $\pi_{at} = p_{1|0}$, $\pi_c = p_{1|1} - p_{1|0} = p_{0|0} - p_{0|1}$ and $\pi_d = 0$. In addition, note that $E[Y^z(0)|at] = E[Y|Z = 0, D = 1]$ and $E[Y^z(1)|nt] = E[Y|Z = 1, D = 0]$, so they are point identified. Furthermore, under Assumptions 1 and 3, it is possible to construct bounds on $E[Y^z(1)|at]$, $E[Y^z(0)|nt]$, $E[Y^z(0)|c]$ and $E[Y^z(1)|c]$ by employing a trimming procedure similar to that used in Lee (2009) and Zhang et al. (2008) in a different context. For instance, consider constructing bounds for $E[Y^z(0)|nt]$. The average outcome for the individuals in the $(Z, D) = (0, 0)$ group can be written as:

$$E[Y|Z = 0, D = 0] = \frac{\pi_{nt}}{\pi_{nt} + \pi_c} \cdot E[Y^z(0)|nt] + \frac{\pi_c}{\pi_{nt} + \pi_c} \cdot E[Y^z(0)|c]. \quad (9)$$

The proportion of never takers in the observed group $(Z, D) = (0, 0)$ is point identified as $\pi_{nt}/(\pi_{nt} + \pi_c) = p_{0|1}/p_{0|0}$. Thus, $E[Y^z(0)|nt]$ can be bounded from above by the expected value of Y for the $p_{0|1}/p_{0|0}$ fraction of *largest values* of Y for those in the observed group $(Z, D) = (0, 0)$. Similarly, it can be bounded from below by the expected value of Y for the $p_{0|1}/p_{0|0}$ fraction of *smallest values* of Y for those in the same observed group. Following this approach, bounds on $E[Y^z(0)|c]$ can also be constructed from (9), and bounds on $E[Y^z(1)|at]$ and $E[Y^z(1)|c]$ can be derived similarly based on the observed group $(Z, D) = (1, 1)$. Finally, note that the bounds on $E[Y^z(0)|nt]$ and $E[Y^z(1)|at]$ can be used to construct bounds on $LNATE_{nt}$ and $LNATE_{at}$, respectively, as the other term in the definition of each of these $LNATE$s is point identified (see equation (8)).

An important step in deriving bounds for $MATE$ consists of writing it in different ways as a function of terms that are point or partially identified under Assumptions 1 and 3. Under

Assumption 3, $MATE$ in (3) can be written as:

$$\begin{aligned}
& MATE \\
= \; & \pi_c LMATE_c && (10)\\
= \; & \pi_{nt} E\left[Y^z(0)\,|nt\right] + \pi_{at} E\left[Y^z(0)\,|at\right] + \pi_c E\left[Y^z(1)\,|c\right] - \pi_c LNATE_c - E\left[Y^z(0)\right] && (11)\\
= \; & E\left[Y^z(1)\right] - \pi_{at} E\left[Y^z(1)\,|at\right] - \pi_{nt} E\left[Y^z(1)\,|nt\right] - \pi_c E\left[Y(1, D(0))\,|c\right] && (12)\\
= \; & E\left[Y^z(1)\right] - E\left[Y^z(0)\right] - \pi_{at} LNATE_{at} - \pi_{nt} LNATE_{nt} - \pi_c LNATE_c. && (13)
\end{aligned}$$

Each of the equations in (10)-(13) exploits different information in the data and, depending on the additional assumptions imposed below, may generate different bounds on $MATE$. Equation (10) is the simplest form of $MATE$ and uses the fact that $LMATE_{nt} = LMATE_{at} = 0$. Equation (11) exploits the fact that $E\left[Y^z(0)\right]$ and $E\left[Y^z(0)\,|at\right]$ are point identified by adding and subtracting $E\left[Y^z(0)\,|k\right]$ for $k = nt, at$ to equation (10). It also adds and subtracts $E\left[Y^z(1)\,|c\right]$ to take advantage of the information available in the data about it and of some assumptions on $LNATE_c$ to be considered below. Equation (12) adds and subtracts $E\left[Y^z(1)\,|k\right]$ for $k = nt, at$ to (10) to exploit the point identification of $E\left[Y^z(1)\right]$ and $E\left[Y^z(1)\,|at\right]$. The last equation uses the fact that $MATE = ATE_{ZY} - NATE$. It exploits point identification of ATE_{ZY}, $E\left[Y^z(0)\,|at\right]$ and $E\left[Y^z(1)\,|nt\right]$, as well as information about $E\left[Y^z(1)\,|c\right]$ and some assumptions on the $LNATE$s to be considered below.[5] Note that it is not possible to derive bounds for $MATE$ without further assumptions because the data contain no information on the potential outcome $Y(1, D(0))$ for compliers, so the term $E\left[Y(1, D(0))\,|c\right]$ appearing either explicitly or implicitly in equations (10)-(13) is not identified.

We next consider two sets of assumptions relating the unidentified terms in equations (10)-(13) to the point or partially identified terms. The specific approach followed to derive bounds on $MATE$ consists of obtaining bounds for each of the non point-identified terms in equations (10)-(13), plugging them in the corresponding equations above, and then comparing the resulting bounds to rule out lower (upper) bounds that are always smaller (greater) than the others. The first set of assumptions we consider to derive bounds on $MATE$ and $LATE$ involves weak monotonicity of mean potential outcomes within strata.

Assumption 4. (*Weak Monotonicity of Mean Potential Outcomes Within Strata).*

4.1. $E[Y^z(1)\,|c] \geq E[Y(1, D(0))\,|c]$. *4.2.* $E[Y(1, D(0))\,|k] \geq E[Y^z(0)\,|k]$, for $k = nt, at, c$.

Assumption 4.1 implies that $LMATE_c (= LATE) \geq 0$, so that the treatment has a non-negative average effect on the outcome for the compliers. When combined with Assumption 3,

[5] It does not make a difference if we write separately each of the terms in $LNATE_{at}$ and $LNATE_{nt}$ in equations (10)-(13) since, as previously discussed, only one of the terms in each of the effects is point identified. Conversely, it is convenient not to brake up $LNATE_c$ and $LMATE_c$ into each of their terms because none of them is point identified and below we consider assumptions about the signs of $LNATE_c$ and $LMATE_c$.

it also implies that $MATE = \pi_c LMATE_c \geq 0$. Assumption 4.2 states that for each strata, the instrument has a non-negative average effect on the outcome net of the effect that works through the treatment status. It requires that $LNATE \geq 0$ for all strata, which implies that $NATE \geq 0$. Hence, under Assumptions 3 and 4, we have $ATE_{ZY} \geq 0$, and the instrument is assumed to have a non-negative average effect on the outcome. We note that although assuming that $LNATE_{nt}$ and $LNATE_{at}$ are non-negative is not strictly necessary to derive bounds on $MATE$, it is helpful in tightening the bounds. For example, under Assumptions 1, 3 and 4, the upper bound for $E[Y^z(0)|nt]$ is the minimum of the upper bound derived using the trimming procedure described above and $E[Y|Z=1,D=0]$, which comes from Assumption 4.2 since $E[Y(1,D(0))|nt] = E[Y^z(1)|nt] = E[Y|Z=1,D=0]$.

Assumptions similar to those in Assumption 4 have been considered for partial identification of average treatment effects in IV models (e.g., Manski and Pepper, 2000, 2009) and in other settings (Manski, 1997; Cai et al., 2008; Sjölander, 2009). For instance, Manski and Pepper (2000, 2009) consider the "monotone treatment response" (MTR) assumption, which states that the potential outcome is a monotone function of the treatment, or $Y_i(1) \geq Y_i(0)$ for all i. In contrast to the MTR assumption, note that Assumption 4.1 allows some individual effects of the treatment on the outcome to be negative by imposing this monotone restriction on the mean potential outcomes for the compliers.

Let y_r^{zd} be the r-th quantile of Y conditional on $Z=z$ and $D=d$. For ease of exposition, suppose Y is continuous so that $y_r^{zd} = F^{-1}_{Y|Z=z,D=d}(r)$, with $F_{\cdot}(\cdot)$ the conditional density of Y given $Z=z$ and $D=d$.[6] We denote by $U^{z,k}$ and $L^{z,k}$ the upper and lower bounds, respectively, on the mean potential outcome $Y^z(z)$ for the strata k derived using the trimming procedure described above, where $z = \{0,1\}$ and $k = \{at, nt, c\}$. The following proposition presents bounds on $LATE$ under Assumptions 1 through 4.

Proposition 2 *If Assumptions 1 through 4 hold, then*

$$0 \leq LATE \leq \frac{\min\left\{U^1, U^2, U^3, U^4\right\}}{E\left[D|Z=1\right] - E\left[D|Z=0\right]},$$

[6]It is straightforward to adapt the trimming procedure above and hence the bounds presented in this section to settings where the outcome has discrete support (see, e.g., Lee 2009, footnote 27).

where

$$\begin{aligned}
U^1 &= \left(p_{1|1} - p_{1|0}\right)\left(U^{1,c} - L^{0,c}\right) \\
U^2 &= p_{0|1} \min\left\{E\left[Y|Z=1, D=0\right], U^{0,nt}\right\} + p_{1|0} E\left[Y|Z=0, D=1\right] \\
&\quad + \left(p_{1|1} - p_{1|0}\right) U^{1,c} - E\left[Y|Z=0\right] \\
U^3 &= E\left[Y|Z=1\right] - p_{1|0} \max\{E\left[Y|Z=0, D=1\right], L^{1,at}\} - p_{0|1} E\left[Y|Z=1, D=0\right] \\
&\quad - \left(p_{1|1} - p_{1|0}\right) L^{0,c} \\
U^4 &= E\left[Y|Z=1\right] - E\left[Y|Z=0\right] - p_{1|0} \max\left\{0, L^{1,at} - E[Y|Z=0, D=1]\right\} \\
&\quad - p_{0|1} \max\left\{0, E[Y|D=1, D=0] - U^{0,nt}\right\} \\
U^{0,nt} &= E[Y|Z=0, D=0, Y \geq y^{00}_{1-(p_{0|1}/p_{0|0})}] \\
U^{1,c} &= E[Y|Z=1, D=1, Y \geq y^{11}_{(p_{1|0}/p_{1|1})}] \\
L^{1,at} &= E[Y|Z=1, D=1, Y \leq y^{11}_{(p_{1|0}/p_{1|1})}] \\
L^{0,c} &= E[Y|Z=0, D=0, Y \leq y^{00}_{1-(p_{0|1}/p_{0|0})}].
\end{aligned}$$

Proof. See Appendix.

Each of the terms U^1, U^2, U^3 and U^4 in the upper bound for $LATE$ comes from the corresponding equation in (10)-(13), while the lower bound of zero comes directly from Assumption 4.1 and equation (10). For example, U^4 comes from equation (13). The first two terms of U^4 identify the average effect of Z on Y, the third term is the lower bound on $LNATE_{at}$ times $p_{1|0}$ (recall $p_{1|0} = \pi_{at}$), the fourth term is the lower bound on $LNATE_{nt}$ times $p_{0|1}$ (recall $p_{0|1} = \pi_{nt}$), and the lower bound on $LNATE_c$ is zero. Note that the bounds in Proposition 2 imply that under Assumptions 1 to 4 the upper bound on $LATE$ is at most equal to the IV estimator in (2), since $U^4 \leq E[Y|Z=1] - E[Y|Z=0]$.

In contrast to Assumption 4, the second set of assumptions we consider does not impose restrictions on the sign of $LATE$. It involves weak monotonicity of mean potential outcomes across strata.

Assumption 5. (*Weak Monotonicity of Mean Potential Outcomes Across Strata).*

5.1. $E\left[Y\left(1, D\left(0\right)\right)|c\right] \geq E\left[Y^z\left(1\right)|nt\right]$. *5.2.* $E\left[Y^z\left(1\right)|at\right] \geq E\left[Y\left(1, D\left(0\right)\right)|c\right]$. *5.3.* $E\left[Y^z\left(0\right)|c\right] \geq E\left[Y^z\left(0\right)|nt\right]$. *5.4.* $E\left[Y^z\left(0\right)|at\right] \geq E\left[Y^z\left(0\right)|c\right]$. *5.5.* $E\left[Y^z\left(1\right)|c\right] \geq E\left[Y^z\left(1\right)|nt\right]$. *5.6.* $E\left[Y^z\left(1\right)|at\right] \geq E\left[Y^z\left(1\right)|c\right]$.

Assumption 5 states that the mean potential outcomes of the always takers are greater than or equal to those of the compliers, and that these in turn are greater than or equal to those of the never takers. Assumption 5 formalizes the notion that some strata are likely to have more favorable characteristics and thus better mean potential outcomes than others. For example, in the context of the empirical application presented in Section 4, Assumption 5 states that the mean potential earnings of those who attain a high school, GED, or vocational degree only if assigned to enroll in a training program are greater (less) than or equal to the mean potential earnings of those who never (always) receive a degree whether or not assigned to enroll in training. Two attractive features of Assumption 5 are (1) it may be substantiated with economic theory in practice and (2) it contains testable implications. The combination of Assumptions 1, 3 and 5 imply that $E[Y|Z=0,D=1] \geq E[Y|Z=0,D=0]$ and $E[Y|Z=1,D=1] \geq E[Y|Z=1,D=0]$. These two inequalities follow from equation (9) and the corresponding equation for the observed group $(Z,D)=(1,1)$, respectively. They can be used in practice to falsify the assumptions. Moreover, it is possible to get indirect evidence about the plausibility of Assumption 5 by looking at relevant average baseline characteristics (e.g., pre-treatment outcomes) of the different strata. These tools will be implemented and further discussed in the context of our empirical analysis.

Assumption 5 is related to, but different from, the monotone instrumental variable (MIV) assumption in Manski and Pepper (2000). The MIV assumption states that mean potential outcomes as a function of the treatment vary weakly monotonically across subpopulations defined by specific observed values of the instrument: $E[Y(d)|Z=1] \geq E[Y(d)|Z=0]$ for $d=\{0,1\}$. It relaxes the traditional mean independence assumption in IV models that requires the previous inequality to hold with equality, by requiring the direction of the endogeneity of Z to be known. Assumption 5 differs from the MIV assumption in at least two important ways. First, Assumption 5 refers to potential outcomes that explicitly allow the instrument to have a causal effect on the outcome (through D and other channels) by writing them as a function of the treatment and the instrument. Second, Assumption 5 imposes weak inequality of the different mean potential outcomes across subpopulations defined by specific values of the potential treatment status (principal strata).

Assumptions 5.1 and 5.2 provide a lower and an upper bound for $E[Y(1,D(0))|c]$, respectively. Assumptions 5.3-5.6 are not strictly necessary to derive bounds for $MATE$, but they are helpful in tightening the bounds. For example, combining Assumption 5.3 with equation (9) yields $E[Y|Z=0,D=0] \geq E[Y^z(0)|nt]$, where by definition $E[Y|Z=0,D=0]$ is less than or equal to $U^{0,nt}$, the upper bound for $E[Y^z(0)|nt]$ derived using the trimming procedure described above and formally defined in Proposition 2. The following proposition presents bounds on $LATE$ employing Assumption 5.

Proposition 3 *If Assumptions 1, 2, 3 and 5 hold, then*

$$\frac{\max\{\overline{L}_1, \overline{L}_2\}}{E[D|Z=1]-E[D|Z=0]} \leq LATE \leq \frac{\overline{U}}{E[D|Z=1]-E[D|Z=0]},$$

where

$$
\begin{aligned}
\overline{L}^1 &= -\left(p_{1|1} - p_{1|0}\right) \cdot \left(U^{1,at} - \max\left\{L^{1,c}, E\left[Y|Z=1, D=0\right]\right\}\right) \\
\overline{L}^2 &= -p_{1|1}\left(U^{1,at} - E\left[Y|Z=1, D=1\right]\right) \\
\overline{U} &= \left(p_{1|1} - p_{1|0}\right)\left(E\left[Y|Z=1, D=1\right] - E\left[Y|Z=1, D=0\right]\right) \\
U^{1,at} &= E[Y|Z=1, D=1, Y \geq y^{11}_{1-(p_{1|0}/p_{1|1})}] \\
L^{1,c} &= E[Y|Z=1, D=1, Y \leq y^{11}_{1-(p_{1|0}/p_{1|1})}].
\end{aligned}
$$

Proof. See Appendix.

The lower bounds for $MATE$, $\overline{L}^1$ and $\overline{L}^2$, come from equations (10) and (12), respectively, and they are always greater than or equal to those derived using equations (11) and (13). The upper bound in Proposition 3 comes from the bounds derived using equations (10) and (12), which under the assumptions in Proposition 3 are equal to each other, and are always less than or equal to those based on equations (11) and (13). The fact that none of the bounds in Proposition 3 comes from equations (11) and (13) is intuitive because these two equations exploit assumptions on the sign of the $LNATE$s, which are not imposed in Proposition 3. The lower bound on $LATE$ in Proposition 3 is always less than or equal to zero because $p_{1|1} - p_{1|0} = \pi_c \geq 0$ and $U^{1,at}$ is always greater than or equal to $E\left[Y|Z=1, D=1\right]$, $L^{1,c}$, and $E[Y|Z=1, D=0]$ (from the testable implications discussed above). Thus, the bounds in Proposition 3 cannot be used to rule out a negative $LATE$. Nevertheless, as illustrated in our empirical application, the upper bound on $LATE$ in this proposition can be informative.

Finally, we combine Assumptions 1 through 5 to construct bounds on $LATE$. Combining Assumptions 4 and 5 yields an additional testable implication: $E[Y|Z=1, D=1] \geq E[Y|Z=0, D=0]$.[7] The following proposition presents the bounds on $LATE$ for this case.

Proposition 4. *If Assumptions 1 through 5 hold,*

$$
0 \leq LATE \leq \frac{\min\{\widetilde{U}_1, \widetilde{U}_2\}}{E[D|Z=1] - E[D|Z=0]},
$$

[7]Note that Assumptions 4 and 5 imply $E[Y^z(1)|at] \geq E[Y^z(0)|at] \geq E[Y^z(0)|c] \geq E[Y^z(0)|nt]$ and $E[Y^z(1)|c] \geq E[Y^z(0)|c] \geq E[Y^z(0)|nt]$. The result follows from combining these inequalities with equation (9) and the corresponding equation for the observed group $(Z, D) = (1, 1)$.

where

$$
\begin{aligned}
\widetilde{U}_1 &= E\left[Y|Z=1\right] - p_{1|0}\max\{E\left[Y|Z=1,D=1\right], E\left[Y|Z=0,D=1\right]\} \\
&\quad - \left(p_{1|1}-p_{1|0}\right)\max\{E[Y|Z=1,D=0], E[Y|Z=0,D=0]\} \\
&\quad - p_{0|1}E\left[Y|Z=1,D=0\right] \\
\widetilde{U}_2 &= E\left[Y|Z=1\right] - E\left[Y|Z=0\right] \\
&\quad - p_{1|0}\max\{0, E\left[Y|Z=1,D=1\right] - E\left[Y|Z=0,D=1\right]\} \\
&\quad - p_{0|1}\max\{0, E\left[Y|Z=1,D=0\right] - E\left[Y|Z=0,D=0\right]\} - \left(p_{1|1}-p_{1|0}\right)\cdot \\
&\quad \max\{0, E\left[Y|Z=1,D=0\right] - U^{0,c}, E\left[Y|Z=1,D=0\right] - E\left[Y|Z=0,D=1\right]\} \\
U^{0,c} &= E[Y|Z=0,D=0,Y \geq y^{00}_{(p_{0|1}/p_{0|0})}].
\end{aligned}
$$

Proof. See Appendix.

The upper bound for $MATE$, $\widetilde{U}$, comes from equation (12), while $\widetilde{U}_2$ comes from equation (13). Similar to Proposition 2, $\widetilde{U}_2$ implies that the upper bound on $LATE$ in Proposition 4 is at most equal to the IV estimator in (2), and Assumptions 4.1 and equation (10) imply that the lower bound on $LATE$ is zero.

It is important to note that the particular conditions imposed in Assumptions 4 and 5 can be changed depending on their plausibility, identifying power and the economic theory behind any particular application. First, some particular assumptions can be dropped if they are not plausible or needed in a given application. For instance, as previously mentioned, Assumptions 5.3-5.6 and 4.2 for the *nt* and *at* strata are not strictly necessary to derive bounds on $LATE$. Similarly, some assumptions can be dropped if interest lies only on a lower or upper bound for $LATE$. Second, the direction of the weak inequalities, including that in Assumption 3, can be reversed depending on the empirical setting. Third, some specific potential outcomes can be changed. For instance, Assumption 5.1 could be changed to $E\left[Y\left(1,D\left(0\right)\right)|c\right] \geq E\left[Y^z\left(0\right)|nt\right]$, which may be easier to justify in some empirical settings. Finally, we also note that it is possible to construct bounds for $LATE$ without Assumptions 4 and 5 if we assume that the support of $Y\left(\cdot\right)$ is bounded, so that $E\left[Y\left(1,D\left(0\right)\right)|c\right]$ is also bounded. In any of these instances, the same approach employed here to derive the bounds in Propositions 2 to 4 can be followed to derive bounds on $LATE$.

3 Bounds on the Effect of Attaining a Degree on Labor Market Outcomes

There is a large empirical literature analyzing the effect of education on labor market outcomes (e.g., Card, 1999), as well as the effects of degree attainment upon earnings (e.g., Hungerford and Solon, 1987; Cameron and Heckman, 1993; Jaeger and Page, 1996; Flores-Lagunes and Light, 2010). In this section, we illustrate the identifying power of the bounds presented above by analyzing the effect of attaining a GED, high school, or vocational degree on labor market outcomes using randomization into a training program as an instrument. The program we consider is Job Corps (JC), the largest and most comprehensive job training program for economically disadvantaged youth aged 16 to 24 years old. In addition to academic and vocational training, JC provides its participants a variety of services such as health services, counseling, job search assistance, social skills training and a stipend during program enrollment, as well as room and board for those residing at the JC centers during program enrollment. We concentrate on the estimation of the returns to attaining any combination of GED, high school, or vocational degrees on labor market outcomes because many JC participants attain at least two degrees (a GED or high school degree plus a vocational degree), and thus breaking up the effects of the different degrees would require additional assumptions. As a reference, JC has been found to have impacts on participants' earnings that are roughly equivalent to a year of regular schooling (Schochet et al., 2001; Lee, 2009). Similarly, Flores et al. (2010) find that the estimated impact of the amount of academic and vocational instruction received in JC on earnings implies annualized returns that are similar to those of IV estimates of the returns to an additional year of regular schooling for people with comparably low educational attainment.

We use data from the National Job Corps Study (NJCS), a randomized experiment performed in the mid-1990s to evaluate the effectiveness of JC. A random sample of all pre-screened eligible applicants in the 48 contiguous states and the District of Columbia was randomly assigned into treatment and control groups, with the second group being denied access to JC for three years. Both groups were tracked with a baseline interview immediately after randomization and thereafter at 12, 30 and 48 months.[8] The specific sample we use consists of all individuals with non-missing values on the randomized treatment status, the variables regarding the attainment of a GED, high school or vocational degree, and the outcome variables considered. We focus on the outcomes measured twelve quarters after random assignment, which corresponds to the time the embargo from the program ended for the control group. The treatment and control groups employed consist of 5,045 and 2,975 individuals, respectively.[9]

We use as an instrumental variable the randomized indicator for whether or not the individual was assigned to participate in JC. For simplicity, we also refer to the instrument as the "program status".[10] Table 2 presents point estimates for some relevant quantities. The

[8] For further description of the JC program and the NJCS see Schochet, Burghardt and Glazerman (2001), Lee (2009) and Flores-Lagunes, Gonzalez and Neumann (2010).

[9] In this application we abstract from the problems of sample attrition over time and missing values. Lee (2009), who employs a similar sample, suggests that the attrition/non-response problem is not serious.

[10] There exits non-compliance with the assigned treatment in the NJCS. The proportion of those in the treatment group who enroll in JC was 73 percent, and the proportion of those in the control group that managed to enroll in JC was 1.4 percent. We discuss below the implications of non-compliance with the assigned treatment

ATE of the program status on the probability of being employed twelve quarters after random assignment is 4 percentage points, while the ATE on weekly earnings is \$18.1. The ATE of the program status on the probability of attaining a high school, GED, or vocational degree is 21 percentage points. All three effects are highly statistically significant.

The large effect of the program status on the attainment of a degree suggests that this instrument satisfies Assumption 2 (non-zero first-stage). From Table 2, the IV point estimate for the effect of attaining a degree on employment and weekly earnings twelve months after randomization using program status as an instrument is 19.4 percentage points and \$86.63, respectively, and both are highly statistically significant. These are point estimates of $LATE$ in (2) under Assumptions 1 to 3 plus the exclusion restriction assumption in (1). In this context, the exclusion restriction assumption requires that all the effect of the program status on employment and earnings works through the attainment of a GED, high school, or vocational degree. Nevertheless, this assumption is likely violated in this setting because JC may have an effect on the outcomes through other services, such as job search services or social skills training. In fact, below we provide some evidence that JC has an effect on both outcomes net of the effect that works through the attainment of a degree, which under the assumptions considered implies that the exclusion restriction is violated. Thus, we apply the bounds for $LATE$ derived in this paper to learn about the effect of interest using program status as an invalid instrument.

We start by discussing the assumptions employed in the paper in the context of this empirical application. Assumption 3 (individual-level monotonicity of Z on D) states that being assigned to participate in JC has a non-negative individual-level effect on the attainment of a GED, high school or vocational degree, so that there are no individuals who would obtain such a degree if they were not assigned to participate in JC and would not if they were assigned. This assumption is plausible in this setting given that JC facilitates the obtainment of such a degree. In this context, the never (always) takers are those individuals who would never (always) obtain a degree regardless of whether or not they are assigned to participate in JC, and the compliers are those who would obtain a degree if they were assigned to participate in JC but would not if they were not assigned. Table 2 shows the estimated proportions of each of these strata, along with bootstrap standard errors.

Assumption 4.1 states that the attainment of a degree has a non-negative average effect on employment and earnings for the compliers, which is consistent with conventional human capital theories in economics. Assumption 4.2 states that the $LNATE$ for all strata are non-negative, or that the combination of the rest of the channels through which the program status affects the outcome has a non-negative average effect on labor market outcomes for all strata. This assumption is likely satisfied in this application because the other components of the JC

for the interpretation of our results.

program (e.g., job search assistance, social skills training, health services) also aim to improve the participants' future labor market outcomes.

Assumption 5 states that the average potential outcomes of the compliers are no less (no greater) than the corresponding average potential outcomes of the never (always) takers. We believe this assumption is also likely to hold in our application given the characteristics of the individuals expected to belong in each strata. For instance, we would expect individuals with more (less) favorable traits to succeed in the labor market (e.g., discipline) to belong to the always-taker (never-taker) strata than the complier strata. Thus, even though Assumption 5 is not directly testable, indirect evidence regarding its plausibility can be gained from comparing average baseline characteristics that are closely related to the outcomes (e.g., values of the outcomes prior to treatment) for the different strata. If these comparisons suggest that the compliers have better (worse) average baseline characteristics than the always (never) takers, Assumption 5 is less likely to hold. In the current application, the probability of being employed and the average weekly earnings in the year prior to randomization for both the always takers and the compliers are statistically greater than those of the never takers, while the differences of those two variables between the always takers and the compliers are not statistically different from zero.[11] Hence, the data does not provide indirect evidence against Assumption 5. In addition to analyzing baseline characteristics, one can check the testable implications of Assumptions 4 and 5 discussed in the previous section. The last three rows of Table 2 verify that these testable implications hold in the data, so the assumptions are not falsified by the data.

Table 3 shows the estimated bounds for the employment and earnings outcomes for each of the bounds in Propositions 2 through 4. For completeness, we also present the bounds for the $MATE$ in (3). We provide standard errors for each of the bounds to give a sense of the accuracy with which they are estimated.[12] In general, the bounds in Table 3 are precisely estimated. We abstract from providing valid confidence sets since it is not straightforward to construct them based on the standard errors reported in Table 3.[13]

[11] The probability of being employed in the year prior to randomization for the never takers, compliers and always takers are, respectively, (standard errors in parenthesis) 0.153 (.009); 0.205 (.046); 0.216 (.011). The corresponding numbers for the average weekly earnings in the year prior to randomization are 86.26 (2.57); 117.73 (12.95); 109.74 (3.09). The means for the never and always takers are calculated from the groups with $(Z_i, D_i) = (1, 0)$ and $(Z_i, D_i) = (0, 1)$, respectively. The mean for the compliers is estimated by writing it as a function of the population mean, the means for the never and always takers, and the strata proportions in the population.

[12] The standard errors for the estimators of the bounds not involving minimum or maximum operators are obtained with 5,000 bootstrap replications. For the estimators of bounds involving those two operators, we combine the bootstrap results for the potential bounds not involving those two operators with the results from Clark (1961), who provides an algorithm to approximate the variance of the maximum of two or more random variables having a joint normal distribution. Finally, for those bounds truncated at zero, we follow Cai et al. (2008) and calculate the standard errors for the estimators employing the formula for a truncated (at zero) normal distribution.

[13] Recent work on inference for partially identified models defined by moment inequalities includes Cher-

We begin by focusing on $MATE$, which gives the part of the average effect of the program status on the outcomes that works through the attainment of a degree. These bounds are shown in the first and third row of Table 3 for the employment and earnings outcomes, respectively. Under Proposition 2, the lower bound on $MATE$ equals zero and the estimated upper bound equals the estimated ATE of the program on the outcome. Replacing Assumption 4 with Assumption 5 (Proposition 3) yields an estimated upper bound of 3 percentage points for the employment outcome and \$14.7 for weekly earnings. This implies that relative to the ATE of program status on the probability of employment (earnings), at most 75 (81) percent of the average effect of the program status on the employment (earnings) outcome is due to the attainment of a degree. When both Assumptions 4 and 5 are used (Proposition 4), the bounds on $MATE$ imply that the part of the average effect of program status on employment (earnings) that is due to the obtainment of a degree is at most half (60%). The fact that the estimated upper bounds on $MATE$ under Assumption 5 (Proposition 3) or Assumptions 4 and 5 (Proposition 4) are considerably below the ATE of program status on the outcomes strongly suggests a failure of the exclusion restriction assumption, and thus the likely unreliability of the conventional $LATE$ point estimates in Table 2. Thus, when the researcher is confident about Assumptions 4 and 5 (as we are in this application), our bounds for $LATE$ can be employed to shed light on the plausibility of the exclusion restriction assumption.

Table 3 also shows the estimated bounds for $LATE$ under the different assumptions considered in the paper (rows 2 and 4). In this empirical application, the $LATE$ in Proposition 1 is interpreted as the local average effect for compliers of attaining a GED, high school or vocational degree on the outcome, when assigned to participate in JC. Note that given the imperfect compliance with the random assignment present in the NJCS, $LATE$ cannot be interpreted as the local average effect for compliers when enrolled in JC. Instead, it has an interpretation similar to that of an "intention-to-treat" parameter.[14]

Under Assumptions 1 through 4 (Proposition 2), the lower bound on $LATE$ for both outcomes is zero, and the estimated upper bounds equal the IV point estimates in Table 2 that assume the validity of the instrument. These bounds come directly from Assumption 4, and hence the data do not provide any additional information to sharpen the bounds when combined with those assumptions. It is important to note, however, that these results imply that the IV point estimate provides an upper bound for $LATE$ in (6) under those assumptions. Therefore, knowledge of the direction of the net and mechanism average effects of the instrument on the outcome provides useful information about the conventional IV estimator.

nozhukov et al. (2007), Bugni (2010), Romano and Shaikh (2010) and Andrews and Soares (2010). We leave for future research the derivation of valid confidence sets for the bounds developed herein.

[14] If the parameter of interest is the local average effect for compliers when enrolled in JC, the approach developed here can be extended to bound that parameter while also addressing the non-compliance problem, but that extension is beyond the scope of the present paper.

The bounds obtained by employing Assumption 5 instead of Assumption 4 (Proposition 3) are more informative with respect to the upper bound on $LATE$. They yield an estimated upper bound on $LATE$ of 15.1 percentage points for the probability of being employed, and of \$70.19 for weekly earnings. These upper bounds are below the $LATE$ point estimates in Table 2, suggesting that, under Assumption 5, these point estimates are upward biased and that the exclusion restriction is violated. The last vertical panel of Table 3 shows the bounds when all five assumptions are combined (Proposition 4). In this case, the lower bound on $LATE$ for both outcomes is zero, which comes directly from Assumption 4. The estimated upper bounds for the $LATE$ of attaining a GED, high school or vocational degree on employment and earnings imply that these effects are at most 10 percentage points and \$53.87, respectively. Both upper bounds are well below the $LATE$ point estimates in Table 2, implying that for both outcomes and under Assumptions 4 and 5, the invalidity of the instrument results in point estimates that are severely upward biased. This is consistent with the intuition that the invalidity of the instrument is due to the availability of other services within JC that affect labor market outcomes positively.

We now discuss the estimated bounds relative to other estimates. First, employing our sample of eligible applicants to the JC program, a simple linear regression of each labor market outcome on the indicator of GED, high school or vocational degree attainment yields estimates of 0.13 (.01) and \$63.1 (4.48) for employment and earnings, respectively (standard errors in parentheses). Both are clearly above our preferred estimated upper bounds for $LATE$ under Proposition 4 (0.1 and \$53.87, respectively). Controlling for a set of covariates in these simple regressions brings the estimated degree effect a step closer to the upper bounds under Proposition 4: 0.11 (.01) and \$57.13 (4.54), respectively.[15]

A second comparison is relative to the bounds derived by Manski and Pepper (2000). We estimated two sets of their bounds for each outcome. The first is under their monotone instrumental variable and monotone treatment response assumptions (MIV-MTR), while the other is under their monotone treatment selection and MTR assumptions (MTS-MTR). The MTS assumption specializes the MIV assumption to the case when the IV is the realized treatment, in which case their bounds do not require a bounded outcome. As discussed in the Introduction, while the bounds in Manski and Pepper (2000) are closer in spirit to ours in that they do not require a valid instrument, they bound a different parameter than our bounds do, the $ATE = E[Y(1) - Y(0)]$. Thus, one must keep this in mind when comparing them. For employment, the estimated MIV-MTR lower and upper bounds are 0 (0) and 0.49 (.01), respectively, while those for earnings are 0 (0) and \$870.6 (52.7), respectively.[16] These upper bounds for ATE are above all those presented in Table 3 for $LATE$. The estimated MTS-MTR bounds

[15] The covariates we control for are age, age squared, race, gender, indicators for whether the individual is married, a household head, has children, and three indicators for the size of the city of residence.

[16] These bounds require a bounded outcome. For earnings, we use the in-sample maximum as the upper bound.

on ATE are closer but still somewhat wider than our bounds on $LATE$ under Proposition 4. For employment, the estimated MTS-MTR lower and upper bounds are 0 (0) and 0.13 (.01), respectively, while those for earnings are 0 (0) and \$63.1 (4.48), respectively.[17]

A third comparison is to studies that estimate degree effects. This literature is not as vast as that analyzing the effect of years of schooling (e.g., Card, 1999), making it more difficult to find estimates of parameters comparable to ours (i.e., $LATE$). Three studies that employ actual information on degree attainment—as opposed to inferring it from years of schooling completed—are Jaeger and Page (1996), Flores-Lagunes and Light (2010) and Cameron and Heckman (1993). The main specification in all three employs OLS on a log-hourly wage model with several control variables. Using CPS data, Jaeger and Page (1996) estimate a 12 percent return to a high school degree for white males (conditional on years of schooling completed) and a 8 percent return to an "occupational associate's degree". The second study employs NLSY79 data and estimates a 20 percent return to a high school or GED degree. Cameron and Heckman (1993) also employ NLSY79 data and report an estimate of 15 percent for high school and 6.2 percent for GED degrees. They also present a specification that controls parametrically for selection, resulting in estimates of 11 percent and 3 percent for high school and GED, respectively. Our preferred estimated upper bound for earnings is \$53.87 which, relative to the average earnings of those individuals assigned to the control group (\$171), represents a return of 31 percent by quarter twelve after random assignment. Hence, the estimates of the above studies fall within the bounds of the $LATE$ in our empirical application.[18]

As a final comparison, we relate our results to the empirical literature on the returns to years of schooling. The average number of actual hours of academic and vocational instruction received while enrolled in JC for those individuals who participated and obtained a degree is 1,448.[19] Considering that a typical high school student receives the equivalent of 1,080 hours of instruction during a school year (Schochet et al., 2001), obtaining a degree is equivalent to about 1.34 years of schooling. Thus, our results above imply an upper bound on the average return to a year of schooling of about 23 percent (31%/1.34). Card (1999) surveyed estimates of the return to a single year of schooling based on IV estimates that exploit institutional features of school systems, which estimate the effect for individuals who would otherwise have relatively low educational attainment and hence are comparable to our sample and parameter

[17] Note that with a binary treatment the upper bound for ATE under the MTS-MTR assumptions is $E[Y|D = 1] - E[Y|D = 0]$.

[18] Indeed, a straightforward comparison is difficult given the differences in (i) samples, (ii) the definition of the outcome variable (earnings versus log-wages) and (iii) the parameters being estimated (ATE versus $LATE$). Nevertheless, the estimates from those papers are a useful point of reference given the apparent lack of studies estimating degree effects employing instrumental variables.

[19] Ideally, we would like an estimate of the average number of actual hours of academic and vocational instruction received by those who were not assigned to participate in JC but completed a degree outside of JC (always takers) in order to have a better estimate of that number for the compliers. Unfortunately, that information is not available from the data.

of interest ($LATE$). The IV estimates he presents range from 6 to 15.3 percent (Table 4 in Card, 1999), and therefore fall within the estimated bounds for the $LATE$ considered in our empirical application.

4 Conclusion

This paper derived nonparametric bounds for local average treatment effects employing an invalid instrument and allowing the outcome to have an unbounded support. We substitute the exclusion restriction assumption in Imbens and Angrist (1994) and Angrist, Imbens and Rubin(1996) with assumptions requiring weak monotonicity of potential outcomes within or across the three principal strata: always takers, never takers and compliers. Our bounds on the effect of attaining a degree on labor market outcomes illustrate their identifying power: they indicate that the local effect when assigned to training for those whose degree attainment is affected by the instrument (random assignment to a training program) is at most 10 percentage points on employment and $54 on weekly earnings. In addition to estimating bounds on the local average treatment effect, the results herein are also potentially useful in the design of experiments. While in some instances it may be difficult or even impossible to randomize a treatment of interest, it may be possible to randomize a variable that affects that treatment. In this case, our approach provides a methodology to obtain bounds on the effect of interest when the randomized variable does not satisfy the exclusion restriction and thus it cannot be used as a valid instrumental variable.

Several extensions of the results in this paper are of interest. First, one could consider settings in which the instrument, the treatment of interest or both are multivalued. In such cases, the number of strata (and thus the number of unidentified objects) increase, so we may expect the bounds to be less informative than in the setting analyzed in this paper. Extensions in this direction can be based on work extending the $LATE$ model in Angrist, Imbens and Rubin(1996) to the multivalued case (Angrist and Imbens, 1995; Nekipelov, 2007). Second, in some applications, the instrument may not be randomly assigned and it may be necessary to adjust for covariates. One could combine the ideas in this paper with work allowing estimation of $LATE$ when the instrument is assumed to be random conditional on a set of covariates (Hirano et al., 2000; Abadie, 2003; Frölich, 2007; Hong and Nekipelov, 2007).[20] While beyond the scope of the current paper, these extensions are at the top of our research agenda.

[20]Note that oftentimes the conditioning on covariates is done to justify the exclusion restriction assumption, which in our approach is not required.

References

[1] Abadie, A. (2003), "Semiparametric Instrumental Variable Estimation of Treatment Response Models", *Journal of Econometrics*, 113(2), 231-263.

[2] Andrews, D. and Soares, G. (2010), "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection", *Econometrica*, 78(1), 119-57.

[3] Angrist, J. (1990), "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records", *American Economic Review*, 80, 313–335.

[4] Angrist, J., and Imbens, G. (1995), "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity", *Journal of the American Statistical Association*, 90, 430–442.

[5] Angrist, J., Imbens, G., and Rubin, D. (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 91, 444–472.

[6] Angrist, J., and Pischke, J-S. (2010), "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics", *Journal of Economic Perspectives,* 24 (2), 3-30.

[7] Balke, A., and Pearl, J. (1997), "Bounds on Treatment Effects from Studies with Imperfect Compliance", *Journal of the American Statistical Association*, 92, 1171–1176.

[8] Bhattacharya, J., Shaikh, A. M. and Vytlacil, E. (2008), "Treatment Effect Bounds under Monotonicity Assumptions: An Application to Swan-Ganz Catheterization", *American Economic Review: Papers and Proceedings*, 98(2), 351–356.

[9] Bugni, F. (2010), "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set", *Econometrica,* 78, 735-53.

[10] Cai, Z., Kuroki, M., Pearl, J. and Tian, J. (2008), "Bounds on Direct Effects in the Presence of Confounded Intermediate Variables", *Biometrics*, 64, 695-701.

[11] Cameron, S., and Heckman, J. (1993), "The Nonequivalence of High School Equivalents", *Journal of Labor Economics*, 11, 1–47.

[12] Card, D. (1999), "The Causal Effect of Education on Earnings" in O.Ashenfelter and D. Card (Eds.) *Handbook of Labor Economics*, Vol. 3, (pp. 1801–1863). Elsevier Science North Holland.

[13] Chernozhukov, V., Hong, H. and Tamer, E. (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models", *Econometrica*, 75, 1243-1284.

[14] Clark, C. (1961), "The Greatest of a Finite Set of Random Variable", *Operations Research*, 145-162.

[15] Conley, T., Hansen, C. and Rossi, P. (2008), "Plausibly Exogenous", mimeo, Graduate School of Business, University of Chicago.

[16] Deaton, A. (2010a), "Instruments, Randomization, and Learning about Development", *Journal of Economic Literature,* 48, 424-455.

[17] Deaton, A. (2010b), "Understanding the Mechanisms of Economic Development", NBER Working Paper 15891.

[18] Duflo, E., Glennerster, R., and Kremer, M. (2008), "Using Randomization in Development Economics Research: a Toolkit", in T.P. Schultz and J. Strauss (Eds.) *Handbook of Development Economics*, Vol. 4, (pp. 3895–3962). Elsevier Science North Holland

[19] Flores, C. and Flores-Lagunes, A. (2010), "Nonparametric Partial Identification of Causal Net and Mechanism Average Treatment Effects", mimeo, Department of Economics, University of Miami.

[20] Flores, C., Flores-Lagunes, A., Gonzalez, A., and Neumann, T. (2010), "Estimating the Effects of Length of Exposure to Instruction in a Training Program: The Case of Job Corps", forthcoming, *Review of Economics and Statistics*.

[21] 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, 651-67.

[22] Flores-Lagunes, A., and Light, A. (2010), "Interpreting Degree Effects in the Returns to Education", *Journal of Human Resources*, 45, 439-467.

[23] Frangakis,C.E. and Rubin D. (2002) "Principal Stratification in Causal Inference", *Biometrics*, 58, 21-29.

[24] Frölich, M. (2007), "Nonparametric IV Estimation of Local Average Treatment Effects with Covariates" *Journal of Econometrics*, 139, 35-75.

[25] Heckman, J., LaLonde, R. and Smith, J. (1999) "The Economics and Econometrics of Active Labor Market Programs" in O. Ashenfelter and D. Card (Eds.) *Handbook of Labor Economics*. (pp. 1865-2097). Elsevier Science North Holland.

[26] Heckman, J. and Urzua, S. (2010), "Comparing IV with Structural Models: What Simple IV can and cannot Identify", *Journal of Econometrics*, 156, 27-37.

[27] Heckman, J., and Vytlacil, E. (1999), "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects", *Proceedings of the National Academy of Sciences*, 96, 4730-4734.

[28] Heckman, J. and Vytlacil, E. (2000), "Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect", NBER Technical Working Paper 259.

[29] Hirano, K., Imbens, G., Rubin, D., and Zhou, X. (2000), "Assessing the Effect of an Influenza Vaccine in an Encouragement Design with Covariates", 2000, *Biostatistics*, 1, 69-88.

[30] Hong, H. and Nekipelov, D. (2007), "Semiparametric Efficiency in Nonlinear LATE Models", mimeo, University of California, Berkeley.

[31] Hungerford, T., and Solon, G. (1987), "Sheepskin Effects in the Returns to Education", *Review of Economics and Statistics*, 69, 175-177.

[32] Imai, K., Keele, L., and Yamamoto, T. (2010), "Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects", Working Paper, Department of Politics, Princeton University.

[33] Imbens, G. (2010) "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)", *Journal of Economic Literature,* 48, 399-423.

[34] Imbens, G. W. and Angrist, J. D. (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, 62 (2), 467-475.

[35] Imbens, G. and Wooldridge, J. (2009), "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 47, 5-86.

[36] Jaeger, D., and Page, M. (1996), "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education", *Review of Economics and Statistics*, 78, 733-740.

[37] Lee, D. (2009), "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects", *Review of Economic Studies*, 76(3), 1071-102.

[38] Manski, C. (1990), "Nonparametric Bounds on Treatment Effects", *American Economic Review: Papers and Proceedings*, 80, 319-323.

[39] Manski, C. (1994), "The Selection Problem", in C. Sims et al. (Eds), *Advances in Econometrics: Sixth World Congress*, (pp. 143-170). Cambridge University Press.

[40] Manski, C. (1997), "Monotone Treatment Response", *Econometrica*, 65, 1311-1334.

[41] Manski, C. and Pepper, J. (2000), "Monotone Instrumental Variables: With an Application to the Returns to Schooling", *Econometrica*, 68, 997-1010.

[42] Manski, C. and Pepper, J. (2009), "More on Monotone Instrumental Variables ", *Econometrics Journal*, 12, S200-S216.

[43] Nekipelov, D. (2007), "Endogenous Multi-Valued Treatment Effect Model Under Monotonicity", mimeo, University of California, Berkeley.

[44] Nevo, A. and Rosen, A. (2008), "Identification with Imperfect Instruments", cemmap working paper CWP16/08.

[45] Oreopoulos, P. (2006), "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter", *American Economic Review*, 96, 152-175.

[46] Pearl, J. (2001), "Direct and Indirect Effects", *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, San Francisco, CA: Morgan Kaufmann, 411-20.

[47] Robins, J. M. (2003), "Semantics of Causal DAG Models and the Identification of Direct and Indirect Effects", in *Highly Structured Stochastic Systems* (Eds., P.J. Green, N.L. Hjort, and S. Richardson), (pp. 70-81). Oxford University Press.

[48] Robins, J. and Greenland, S. (1992), "Identifiability and Exchangeability for Direct and Indirect Effects", *Epidemiology*, 3, 143-155.

[49] Romano, J. and Shaikh, A. (2010), "Inference for the Identified Set in Partially Identified Econometric Models", *Econometrica*, 78 (1), 169-211.

[50] Rosenzweig, M. and Wolpin, K. (1980), "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment", *Econometrica*, 48, 227-240.

[51] Schochet, P., Burghardt, J., and Glazerman, S. (2001), *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes*. Princeton, NJ: Mathematica Policy Research, Inc.

[52] Shaikh, A. M. and Vytlacil, E. (2005), "Threshold Crossing Models and Bounds on Treatment Effects: A Nonparametric Analysis", NBER Technical Working Paper 307.

[53] Sjölander, A. (2009), "Bounds on Natural Direct Effects in the Presence of Confounded Intermediate Variables", *Statistics in Medicine*, 28, 558-71.

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

[55] Zhang, J.L., 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 Science North Holland.

5 Appendix

From Section 2.2, the relevant point identified objects in our setting are: $\pi_{nt} = p_{0|1}$, $\pi_{at} = p_{1|0}$, $\pi_c = p_{1|1} - p_{1|0} = p_{0|0} - p_{0|1}$, $E[Y^z(1)] = E[Y|Z=1]$, $E[Y^z(0)] = E[Y|Z=0]$, $E[Y^z(1)|nt] = E[Y|Z=1, D=0]$, $E[Y^z(0)|at] = E[Y|Z=0, D=1]$, $\pi_{nt}E[Y_i^z(0)|nt] + \pi_c E[Y_i^z(0)|c] = p_{0|0}E[Y_i|Z_i=0, D_i=0]$ and $\pi_{at}E[Y_i^z(1)|at] + \pi_c E[Y_i^z(1)|c] = p_{1|1}E[Y_i|Z_i=1, D_i=1]$.

The trimming-based bounds on mean potential outcomes at the strata level discussed in Section 2.2 are given by: $L^{0,nt} \le E[Y^z(0)|nt] \le U^{0,nt}$; $L^{1,at} \le E[Y^z(1)|at] \le U^{1,at}$; $L^{0,c} \le E[Y^z(0)|c] \le U^{0,c}$ and $L^{1,c} \le E[Y^z(1)|c] \le U^{1,c}$, where

$$\begin{aligned}
L^{0,nt} &= E[Y|Z=0, D=0, Y \le y^{00}_{(p_{0|1}/p_{0|0})}],\ U^{0,nt} = E[Y|Z=0, D=0, Y \ge y^{00}_{1-(p_{0|1}/p_{0|0})}],\\
L^{1,at} &= E[Y|Z=1, D=1, Y \le y^{11}_{(p_{1|0}/p_{1|1})}],\ U^{1,at} = E[Y|Z=1, D=1, Y \ge y^{11}_{1-(p_{1|0}/p_{1|1})}],\\
L^{0,c} &= E[Y|Z=0, D=0, Y \le y^{00}_{1-(p_{0|1}/p_{0|0})}],\ U^{0,c} = E[Y|Z=0, D=0, Y \ge y^{00}_{(p_{0|1}/p_{0|0})}],\\
L^{1,c} &= E[Y|Z=1, D=1, Y \le y^{11}_{1-(p_{1|0}/p_{1|1})}],\ U^{1,c} = E[Y|Z=1, D=1, Y \ge y^{11}_{(p_{1|0}/p_{1|1})}].
\end{aligned}$$

Proof of Proposition 2. We start by deriving bounds for the non-point identified mean potential outcomes of the stratas, and for all the local net and mechanism average treatment effects.

Bounds for $E[Y^z(0)|nt]$: Ass. 4.2 implies $E[Y^z(1)|nt] = E[Y|Z=1, D=0] \ge E[Y^z(0)|nt]$. Ass. 4 does not provide any additional information for a lower bound of $E[Y^z(0)|nt]$. Since $U^{0,nt}$ can be above or below $E[Y|Z=1, D=0]$, we have: $L^{0,nt} \le E[Y^z(0)|nt] \le \min\{U^{0,nt}, E[Y|Z=1, D=0]\}$.[21]

Bounds for $E[Y^z(1)|at]$: Ass. 4.2 implies $E[Y^z(1)|at] \ge E[Y^z(0)|at] = E[Y|Z=0, D=1]$. Ass. 4 does not provide any additional information for an upper bound of $E[Y^z(1)|at]$. Thus we have: $\max\{L^{1,at}, E[Y|Z=0, D=1]\} \le E[Y^z(1)|at] \le U^{1,at}$.

Bounds for $E[Y^z(0)|c]$: Ass. 4.1 and 4.2 imply $E[Y^z(1)|c] \ge E[Y^z(0)|c]$, which implies that $U^{1,c}$ is another upper bound for $E[Y^z(0)|c]$. Ass. 4 does not provide any additional information for a lower bound of $E[Y^z(0)|c]$. Hence, $L^{0,c} \le E[Y^z(0)|c] \le \min\{U^{0,c}, U^{1,c}\}$.

Bounds for $E[Y^z(1)|c]$: Ass. 4 implies $E[Y^z(1)|c] \ge E[Y^z(0)|c]$, which implies that $L^{0,c}$ is another lower bound for $E[Y^z(1)|c]$. Hence, $\max\{L^{0,c}, L^{1,c}\} \le E[Y^z(1)|c] \le U^{1,c}$.

Bounds for $E[Y(1, D(0))|c]$: Ass. 4.1 and 4.2 imply $E[Y^z(1)|c] \ge E[Y(1, D(0))|c] \ge E[Y^z(0)|c]$, which combined with the results above gives $L^{0,c} \le E[Y(1, D(0))|c] \le U^{1,c}$.

Bounds for $LNATE_{nt}$: From (8), $LNATE_{nt} = E[Y|Z=1, D=0] - E[Y^z(0)|nt]$. Using the bounds previously derived for $E[Y^z(0)|nt]$, and letting $L^{nt} = E[Y|Z=1, D=0] - U^{0,nt}$ and $U^{nt} = E[Y|Z=1, D=0] - L^{0,nt}$, we have:[22] $\max\{0, L^{nt}\} \le LNATE_{nt} \le U^{nt}$.

[21] For brevity, in what follows we omit explicitly specifying when some quantities can be greater or lower than others unless we believe it is necessary. Hence, when min (or max) operators are present, it implies that none of the terms inside them is always lower (greater) than the other(s).

[22] The following equalities are helpful for the rest of the proofs. For scalars a, b, c and d we have: (i) $a -$

Bounds for $LNATE_{at}$: From (8), $LNATE_{at} = E[Y^z(1)|at] - E[Y|Z=0, D=1]$. Using the bounds previously derived for $E[Y^z(1)|at]$, and letting $L^{at} = L^{1,at} - E[Y|Z=0, D=1]$ and $U^{at} = U^{1,at} - E[Y|Z=0, D=1]$, we have: $\max\{0, L^{at}\} \leq LNATE_{at} \leq U^{at}$.

Bounds for $LNATE_c$: From (8), $LNATE_c = E[Y(1, D(0))|c] - E[Y^z(0)|c]$. Ass. 4.2 directly implies $LNATE_c \geq 0$. Using the bounds previously obtained for the components in $LNATE_c$ we obtain two additional lower bounds: $L^{0,c} - U^{0,c}$ and $L^{0,c} - U^{1,c}$. By definition, $L^{0,c} - U^{0,c} \leq 0$. Also, employing Ass. 4 we have $U^{1,c} \geq E[Y^z(1)|c] \geq E[Y^z(0)|c] \geq L^{0,c}$, so $L^{0,c} - U^{1,c} \leq 0$. Hence, the lower bound for $LNATE_c$ is 0. Using the bounds previously derived for the components of $LNATE_c$, we have the upper bound is $U^{1,c} - L^{0,c}$. Thus, $0 \leq LNATE_c \leq (U^{1,c} - L^{0,c})$.

Bounds for $LMATE_c$: $LMATE_c = E[Y^z(1)|c] - E[Y(1, D(0))|c]$. Ass. 4.1 directly implies $LMATE_c \geq 0$. Using the bounds previously obtained for the components of $LMATE_c$ we obtain two additional lower bounds: $L^{1,c} - U^{1,c}$ and $L^{0,c} - U^{1,c}$. Since $L^{1,c} - U^{1,c} \leq 0$ (by definition) and $L^{0,c} - U^{1,c} \leq 0$ (from above), the lower bound for $LMATE_c$ is 0. Using the bounds previously derived for the components of $LMATE_c$, we have the upper bound is $U^{1,c} - L^{0,c}$. Thus, $0 \leq LMATE_c \leq (U^{1,c} - L^{0,c})$.

We now derive the bounds for $MATE$, starting with the lower bound. We use equations (10) to (13) to derive potential lower bounds for $MATE$ by plugging in the appropriate bounds derived above into the terms that are not point identified. The corresponding four potential lower bounds are:

$$
\begin{aligned}
\Delta_1 &= 0 \\
\Delta_2 &= p_{0|1} L^{0,nt} + p_{1|0} E[Y|Z=0, D=1] + (p_{1|1} - p_{1|0}) \max\{L^{0,c}, L^{1,c}\} \\
&\quad - (p_{1|1} - p_{1|0}) [U^{1,c} - L^{0,c}] - E[Y|Z=0] \\
\Delta_3 &= E[Y|Z=1] - p_{1|0} U^{1,at} - p_{0|1} E[Y|Z=1, D=0] - (p_{1|1} - p_{1|0}) U^{1,c} \\
\Delta_4 &= E[Y|Z=1] - E[Y|Z=0] - p_{1|0} U^{at} - p_{0|1} U^{nt} - (p_{1|1} - p_{1|0}) [U^{1,c} - L^{0,c}].
\end{aligned}
$$

After some algebra, we have $\Delta_2 = -\pi_c(U^{1,c} - \max\{L^{0,c}, L^{1,c}\}) - \pi_{nt}(E[Y^z(0)|nt] - L^{0,nt}) - \pi_c(E[Y^z(0)|c] - L^{0,c})$; $\Delta_3 = -\pi_{at}(U^{1,at} - E[Y^z(1)|at]) - \pi_c(U^{1,c} - E[Y^z(1)|c])$ and $\Delta_4 = \Delta_3 - \pi_{nt}(E[Y^z(0)|nt] - L^{0,nt}) - \pi_c(E[Y^z(0)|c] - L^{0,c})$. Using the fact that $U^{1,c} \geq \max\{L^{0,c}, L^{1,c}\}$ (from above), we have: $\Delta_2 \leq 0$, $\Delta_3 \leq 0$ and $\Delta_4 \leq 0$. Hence, the lower bound for $MATE$ is $\Delta_1 = 0$.

Now consider the upper bound for $MATE$. Plugging in the bounds derived above for the corresponding non-point identified terms in equations (10) to (13) yields the lower bounds U^1, U^2, U^3 and U^4, as given in Prop. 2. After some algebra, we can write: $U^4 - U^2 = p_{1|0}(U^{1,c}-$

$\max\{c,d\} = \min\{a-c, a-d\}$; (ii) $a - \min\{c,d\} = \max\{a-c, a-d\}$; (iii) $\max\{a,b\} - c = \max\{a-c, b-c\}$; (iv) $\min\{a,b\} - c = \min\{a-c, b-c\}$; (v) $\max\{a,b\} - \min\{c,d\} = \max\{a-c, a-d, b-c, b-d\}$; (vi) $\min\{a,b\} - \max\{c,d\} = \min\{a-c, a-d, b-c, b-d\}$.

$E[Y|Z=0,D=1] - \max\{0, L^{1,at} - E[Y|Z=0,D=1]\}) - p_{1|1}(U^{1,c} - E[Y|Z=1,D=1])$; $U^4 - U^3 = p_{0|1}(E[Y|Z=1,D=0] - L^{0,c} - \max\{0, E[Y|Z=1,D=0] - U^{0,nt}\}) - p_{0|0}(E[Y|Z=0,D=0] - L^{0,c})$; $U^4 - U^1 = (U^4 - U^2) + (U^4 - U^3)$; $U^2 - U^1 = (U^4 - U^3)$; $U^3 - U^1 = (U^4 - U^2)$; and $U^3 - U^2 = (U^4 - U^2) + (U^3 - U^4)$. All six comparisons can be greater or less than zero depending on the data, so no potential upper bound is dropped. To show this, it is enough to get for each comparison one case where the difference can be greater or less than zero. For instance, consider the first difference. $U^4 - U^2$ can be greater or less than zero if $\max\{0, L^{1,at} - E[Y|Z=0,D=1]\} = 0$ and $E[Y|Z=1,D=1] \geq E[Y|Z=0,D=1]$, since $p_{1|1} \geq p_{1|0} \geq 0$ and $(U^{1,c} - E[Y|Z=0,D=1]) \geq (U^{1,c} - E[Y|Z=1,D=1]) \geq 0$. Similar arguments can be made for the rest of the comparisons.

Finally, the bounds for $LATE$ follow directly from the bounds for $MATE$ and the result in Proposition 1. Q.D.E.

Proof of Proposition 3. As before, we first derive bounds for the non-point identified mean potential outcomes of the stratas, and for all the local net and mechanism average treatment effects.

Bounds for $E[Y^z(0)|nt]$: Ass. 5.3 and equation (9) imply $E[Y^z(0)|nt] \leq E[Y|Z=0,D=0]$. Since by definition $E[Y|Z=0,D=0] \leq U^{0,nt}$, the upper bound in this case is $E[Y|Z=0,D=0]$. Ass. 5 does not provide any additional information for a lower bound of $E[Y^z(0)|nt]$. Thus, $L^{0,nt} \leq E\left[Y^z(0)|nt\right] \leq E\left[Y|Z=0,D=0\right]$.

Bounds for $E[Y^z(1)|at]$: Ass. 5.1 and 5.2 imply $E[Y^z(1)|at] \geq E[Y^z(1)|nt] = E[Y|Z=1,D=0]$. Ass. 5.6 and equation (9) for the group $(Z,D)=(1,1)$ yield $E[Y^z(1)|at] \geq E[Y|Z=1,D=1]$.[23] By definition, $E[Y|Z=1,D=1] \geq L^{1,at}$, and note also that $E[Y|Z=1,D=1] \geq E[Y|Z=1,D=0]$. Since Ass. 5 does not provide any additional information for an upper bound of $E[Y^z(1)|at]$, we have $E[Y|Z=1,D=1] \leq E[Y^z(1)|at] \leq U^{1,at}$.

Bounds for $E[Y^z(0)|c]$: Ass. 5.3 and equation (9) yield $E[Y^z(0)|c] \geq E[Y|Z=0,D=0]$, where by definition $E[Y|Z=0,D=0] \geq L^{0,c}$. As for the upper bound, Ass. 5.4 implies $E[Y^z(0)|c] \leq E[Y^z(0)|at] = E[Y|Z=0,D=1]$, which can be greater or less than $U^{0,c}$. Thus, $E[Y|Z=0,D=0] \leq E[Y^z(0)|c] \leq \min\{U^{0,c}, E[Y|Z=0,D=1]\}$.

Bounds for $E\left[Y^z(1)|c\right]$: Ass. 5.6 and equation (9) for the group $(Z,D)=(1,1)$ yield $E[Y^z(1)|c] \leq E[Y|Z=1,D=1]$, where by definition $U^{1,c} \geq E[Y|Z=1,D=1]$. As for the lower bound, Ass. 5.5 implies $E[Y^z(1)|c] \geq E[Y^z(1)|nt] = E[Y|Z=1,D=0]$, which can be greater or less than $L^{1,c}$. Thus, $\max\{L^{1,c}, E[Y|Z=1,D=0]\} \leq E[Y^z(1)|c] \leq E[Y|Z=1,D=1]$.

Bounds for $E[Y(1,D(0))|c]$: Ass. 5.1 implies $E[Y(1,D(0))|c] \geq E[Y^z(1)|nt] = E[Y|Z=1,D=0]$. Combining Ass. 5.2 with the bounds previously derived for $E[Y^z(1)|at]$ yields $E[Y(1,D(0))|c] \leq E[Y^z(1)|at] \leq U^{1,at}$. Hence, $E[Y|Z=1,D=0] \leq E[Y(1,D(0))|c] \leq$

[23] Equation (9) for the group $(Z,D)=(1,1)$ is: $E[Y|Z=1,D=1] = \frac{\pi_{at}}{\pi_{at}+\pi_c} \cdot E[Y^z(1)|at] + \frac{\pi_c}{\pi_{at}+\pi_c} \cdot E[Y^z(1)|c]$.

$U^{1,at}$.

The bounds for $LNATE_{nt}$, $LNATE_{at}$, $LNATE_c$ and $LMATE_c$ follow directly by plugging in the appropriate bounds previously derived for each of their non-point identified components. For instance, for $LNATE_{nt} = E[Y|Z=1, D=0] - E[Y^z(0)|nt]$ we employ the bounds previously derived for $E[Y^z(0)|nt]$ to get $(E[Y|Z=1, D=0] - E[Y|Z=0, D=0]) \leq LNATE_{nt} \leq (E[Y|Z=1, D=0] - L^{0,nt})$.

We now derive the bounds for $MATE$, starting with the lower bound. As before, we use equations (10) to (13) to derive potential lower bounds for $MATE$ by plugging in the appropriate bounds derived above into the terms that are not point identified. The corresponding four potential lower bounds are:

$$\begin{aligned}
\Delta_1 &= \left(p_{1|1} - p_{1|0}\right) \overline{L}_m^c \\
\Delta_2 &= p_{0|1} L^{0,nt} + p_{1|0} E\left[Y|Z=0, D=1\right] + \left(p_{1|1} - p_{1|0}\right) \max\{L^{1,c}, E[Y|Z=1, D=0]\} \\
&\quad - \left(p_{1|1} - p_{1|0}\right) \overline{U}^c - E\left[Y|Z=0\right] \\
\Delta_3 &= E\left[Y|Z=1\right] - p_{1|0} U^{1,at} - p_{0|1} E\left[Y|Z=1, D=0\right] - \left(p_{1|1} - p_{1|0}\right) U^{1,at} \\
\Delta_4 &= E\left[Y|Z=1\right] - E\left[Y|Z=0\right] - p_{1|0} U^{at} - p_{0|1} U^{nt} - \left(p_{1|1} - p_{1|0}\right) \overline{U}^c,
\end{aligned}$$

with $\overline{L}_m^c = \max\left\{L^{1,c}, E\left[Y|Z=1, D=0\right]\right\} - U^{1,at}$, $\overline{U}^c = U^{1,at} - E\left[Y|Z=0, D=0\right]$, and U^{at} and U^{nt} as defined in the proof of Prop. 2. After some algebra we obtain $\Delta_2 - \Delta_1 = \Delta_4 - \Delta_3 = p_{0|1}(L^{0,nt} - E[Y|Z=0, D=0]) \leq 0$, since by definition $L^{0,nt} \leq E[Y|Z=0, D=0]$. We also obtain that $\Delta_3 - \Delta_1 = (p_{1|1} - p_{1|0})(U^{1,at} - \max\{L^{1,c}, E[Y|Z=1, D=0]\}) - p_{1|1}(U^{1,at} - E[Y|Z=1, D=1])$. Note that: (i) $p_{1|1} \geq (p_{1|1} - p_{1|0}) = \pi_c \geq 0$; (ii) $U^{1,at} - E[Y|Z=1, D=1] \geq 0$ by definition; and $(U^{1,at} - \max\{L^{1,c}, E[Y|Z=1, D=0]\}) \geq 0$ since $U^{1,at} \geq E[Y|Z=1, D=1] \geq L^{1,c}$ (by definition) and $U^{1,at} \geq E[Y|Z=1, D=1] \geq E[Y|Z=1, D=0]$; (iii) $(U^{1,at} - \max\{L^{1,c}, E[Y|Z=1, D=0]\}) \geq (U^{1,at} - E[Y|Z=1, D=1])$, since $E[Y|Z=1, D=1] \geq \max\{L^{1,c}, E[Y|Z=1, D=0]\}$ (see part ii). Parts (i) to (iii) imply that $\Delta_3 - \Delta_1$ can be greater or less than zero. After some algebra we have that $\overline{L}^1 = \Delta_1$ and $\overline{L}^2 = \Delta_3$, and thus the lower bound on $MATE$ is $\max\{\overline{L}^1, \overline{L}^2\}$.

Now consider the upper bound for $MATE$. Plugging in the bounds derived above for the corresponding non-point identified terms in equations (10) to (13) yields the following potential upper bounds:

$$\begin{aligned}
\Upsilon_1 &= \left(p_{1|1} - p_{1|0}\right) \overline{U}_m^c \\
\Upsilon_2 &= p_{0|1} E\left[Y|Z=0, D=0\right] + p_{1|0} E\left[Y|Z=0, D=1\right] + \left(p_{1|1} - p_{1|0}\right) E\left[Y|Z=1, D=1\right] \\
&\quad - \left(p_{1|1} - p_{1|0}\right) \overline{L}^c - E\left[Y|Z=0\right] \\
\Upsilon_3 &= E\left[Y|Z=1\right] - p_{1|0} E\left[Y|Z=1, D=1\right] - p_{0|1} E\left[Y|Z=1, D=0\right] \\
&\quad - \left(p_{1|1} - p_{1|0}\right) E\left[Y|Z=1, D=0\right] \\
\Upsilon_4 &= E\left[Y|Z=1\right] - E\left[Y|Z=0\right] - p_{1|0} \overline{L}^{at} - p_{0|1} \overline{L}^{nt} - \left(p_{1|1} - p_{1|0}\right) \overline{L}^c,
\end{aligned}$$

where $\overline{U}^c_m = E\left[Y|Z=1,D=1\right] - E\left[Y|Z=1,D=0\right]$, $\overline{L}^c = E\left[Y|Z=1,D=0\right] - \min\{U^{0,c}, E[Y|Z=0,D=1]\}$, $\overline{L}^{at} = E\left[Y|Z=1,D=1\right] - E\left[Y|Z=0,D=1\right]$, and $\overline{L}^{nt} = E[Y|Z=1,D=0] - E[Y|Z=0,D=0]$. After some algebra we obtain $\Upsilon_2 = \Upsilon_4$, $\Upsilon_3 = \Upsilon_1$ and $\Upsilon_1 - \Upsilon_4 = \pi_c(E[Y|Z=0,D=0] - \min\{U^{0,c}, E[Y|Z=0,D=1]\}) \leq 0$, since $U^{0,c} \geq E[Y|Z=0,D=0]$ and $E[Y|Z=0,D=1] \geq E[Y|Z=0,D=0]$. Thus, the upper bound for $MATE$ equals $\overline{U} \equiv \Upsilon_1$.

Finally, the bounds for $LATE$ follow directly from the bounds for $MATE$ and the result in Proposition 1. Q.D.E.

Proof of Proposition 4. *Bounds for* $E[Y^z(0)|nt]$: Ass. 4.2 implies $E[Y^z(0)|nt] \leq E[Y^z(1)|nt] = E[Y|Z=1,D=0]$; and Ass. 5.3 implies $E[Y^z(0)|nt] \leq E[Y|Z=0,D=0]$ (see proof of Prop. 3), where by definition $U^{0,nt} \geq E[Y|Z=0,D=0]$. Combining the rest of the assumptions does not yield any additional upper bound for $E[Y^z(0)|nt]$ that could be lower than $E[Y|Z=0,D=0]$ or $E[Y|Z=1,D=0]$.[24] Equation (9) and the fact that $E[Y^z(1)|nt] = E[Y|Z=1,D=0]$ imply that $E[Y|Z=1,D=0]$ can be greater or less than $E[Y|Z=0,D=0]$ since, even though $E[Y^z(0)|nt] \leq E[Y^z(1)|nt]$ (by Ass. 4.2), we have that $E[Y^z(1)|nt]$ can be greater or less than $E[Y^z(0)|c]$. Hence, the upper bound for $E[Y^z(0)|nt]$ is $\min\{E[Y|Z=1,D=0], E[Y|Z=0,D=0]\}$. Ass. 4 and 5 do not provide any additional information for a lower bound of $E[Y^z(0)|nt]$. Thus, $L^{0,nt} \leq E\left[Y^z(0)|nt\right] \leq \min\{E[Y|Z=1,D=0], E[Y|Z=0,D=0]\}$.

Bounds for $E[Y^z(1)|at]$: Ass. 4.2 implies $E[Y^z(1)|at] \geq E[Y^z(0)|at] = E[Y|Z=0,D=1]$; and Ass. 5.6 implies $E[Y^z(1)|at] \geq E[Y|Z=1,D=1]$ (see proof of Prop. 3), where by definition $E[Y|Z=1,D=1] \geq L^{1,at}$. Combining Assumptions 4 and 5 does not yield any additional lower bound for $E[Y^z(1)|at]$ that could be greater than $E[Y|Z=1,D=1]$ or $E[Y|Z=0,D=1]$. Equation (9) for the group $(Z,D)=(1,1)$ and the fact that $E[Y^z(0)|at] = E[Y|Z=0,D=1]$ imply that $E[Y|Z=0,D=1]$ can be greater or less than $E[Y|Z=1,D=1]$ since, even though $E[Y^z(0)|at] \leq E[Y^z(1)|at]$ (by Ass. 4.2), we have that $E[Y^z(0)|at]$ can be greater or less than $E[Y^z(1)|c]$. Hence, the lower bound for $E[Y^z(1)|at]$ is $\max\{E[Y|Z=1,D=1], E[Y|Z=0,D=1]\}$. Ass. 4 and 5 do not provide any additional information for an upper bound of $E[Y^z(1)|at]$. Thus, $\max\{E[Y|Z=1,D=1], E[Y|Z=0,D=1]\} \leq E[Y^z(1)|at] \leq U^{1,at}$.

Bounds for $E[Y^z(0)|c]$: Ass. 4 does not provide any information for a lower bound of $E[Y^z(0)|c]$; while Ass. 5.3 and equation (9) yield $E[Y^z(0)|c] \geq E[Y|Z=0,D=0]$, where by definition $E[Y|Z=0,D=0] \geq L^{0,c}$. Regarding an upper bound, the trimming-based bounds imply $E[Y^z(0)|c] \leq U^{0,c}$. Ass. 5.4 implies $E[Y^z(0)|c] \leq E[Y^z(0)|at] = E[Y|Z=0,D=1]$.

[24] For instance, combining Ass. 5.3, 5.4 and 4.2 yields $E[Y^z(1)|at] \geq E[Y^z(0)|at] \geq E[Y^z(0)|c] \geq E[Y^z(0)|nt]$, which implies $E[Y^z(0)|at] = E[Y|T=0,S=1] \geq E[Y^z(0)|nt]$ and $U^{1,at} \geq E[Y^z(1)|at] \geq E[Y^z(0)|nt]$. However, we have that $E[Y|T=0,S=1] \geq E[Y|T=0,S=0]$ and $U^{1,at} \geq E[Y|T=1,S=1] \geq E[Y|T=0,S=0]$.

Finally, Ass. 4 implies $E[Y^z(1)|c] \geq E[Y^z(0)|c]$. Below we show that the upper bound for $E[Y^z(1)|c]$ under Ass. 1, 3, 4 and 5 equals $E[Y|Z=1, D=1]$, so $E[Y^z(0)|c] \leq E[Y|Z=1, D=1]$. Depending on the data, any of the previous three upper bounds for $E[Y^z(0)|c]$ can be less than the other two. Thus, we obtain $E[Y|Z=0, D=0] \leq E[Y^z(0)|c] \leq \min\{U^{0,c}, E[Y|Z=0, D=1], E[Y|Z=1, D=1]\}$.

Bounds for $E[Y^z(1)|c]$: Ass. 4 does not provide any information for an upper bound of $E[Y^z(1)|c]$; while Ass. 5.6 and equation (9) for the group $(Z, D) = (1, 1)$ yield $E[Y^z(1)|c] \leq E[Y|Z=1, D=1]$, where by definition $E[Y|Z=1, D=1] \leq U^{1,c}$. Regarding a lower bound, the trimming-based bounds imply $E[Y^z(1)|c] \geq L^{1,c}$. Ass. 5.5 implies $E[Y^z(1)|c] \geq E[Y^z(1)|nt] = E[Y|Z=1, D=0]$. Finally, Ass. 4 implies $E[Y^z(1)|c] \geq E[Y^z(0)|c]$. Above we showed that the lower bound for $E[Y^z(0)|c]$ under Ass. 1, 3, 4 and 5 equals $E[Y|Z=0, D=0]$, so $E[Y^z(1)|c] \geq E[Y|Z=0, D=0]$. Depending on the data, any of the previous three lower bounds for $E[Y^z(1)|c]$ can be greater than the other two. Thus, we obtain $\max\{L^{1,c}, E[Y|Z=1, D=0], E[Y|Z=0, D=0]\} \leq E[Y^z(1)|c] \leq E[Y|Z=1, D=1]$.

Bounds for $E[Y(1, D(0))|c]$: Ass. 4.2 implies $E[Y(1, D(0))|c] \geq E[Y^z(0)|c]$. From above, the lower bound for $E[Y^z(0)|c]$ under Ass. 1, 3, 4 and 5 equals $E[Y|Z=0, D=0]$. Ass. 5.1 implies $E[Y(1, D(0))|c] \geq E[Y^z(1)|nt] = E[Y|Z=1, D=0]$, which can be greater or less than $E[Y|Z=0, D=0]$ (see above). Hence, $E[Y(1, D(0))|c] \geq \max\{E[Y|Z=0, D=0], E[Y|Z=1, D=0]\}$. Ass. 4.1 implies $E[Y^z(1)|c] \geq E[Y(1, D(0))|c]$. From above, the upper bound for $E[Y^z(1)|c]$ under Ass. 1, 3, 4 and 5 equals $E[Y|Z=1, D=1]$. Note that 5.2 implies $E[Y(1, D(0))|c] \leq E[Y^z(1)|at] \leq U^{1,at}$, but by definition $E[Y|Z=1, D=1] \leq U^{1,at}$. Therefore, $\max\{E[Y|Z=1, D=0], E[Y|Z=0, D=0]\} \leq E[Y(1, D(0))|c] \leq E[Y|Z=1, D=1]$.

Bounds for $LNATE_{nt}$: From (8), $LNATE_{nt} = E[Y|Z=1, D=0] - E[Y^z(0)|nt]$. Using the bounds previously derived for $E[Y^z(0)|nt]$ we have: $\max\{0, \overline{L}^{nt}\} \leq LNATE_{nt} \leq U^{nt}$, with $\overline{L}^{nt}$ and U^{nt} as defined in the proofs of Prop. 3 and 2, respectively.

Bounds for $LNATE_{at}$: From (8), $LNATE_{at} = E[Y^z(1)|at] - E[Y|Z=0, D=1]$. Using the bounds previously derived for $E[Y^z(1)|at]$ we have: $\max\{0, \overline{L}^{at}\} \leq LNATE_{at} \leq U^{at}$, with $\overline{L}^{at}$ and U^{at} as defined in the proofs of Prop. 3 and 2, respectively.

Bounds for $LNATE_c$: From (8), $LNATE_c = E[Y(1, D(0))|c] - E[Y^z(0)|c]$. Ass. 4.2 directly implies $LNATE_c \geq 0$. Using the bounds previously obtained for the components of $LNATE_c$ we obtain six additional potential lower bounds: $E[Y|Z=1, D=0] - U^{0,c}$, $E[Y|Z=1, D=0] - E[Y|Z=1, D=1]$, $E[Y|Z=1, D=0] - E[Y|Z=0, D=1]$, $E[Y|Z=0, D=0] - U^{0,c}$, $E[Y|Z=0, D=0] - E[Y|Z=1, D=1]$ and $E[Y|Z=0, D=0] - E[Y|Z=0, D=1]$. Note that: $E[Y|Z=1, D=0] - E[Y|Z=1, D=1] \leq 0$; $E[Y|Z=0, D=0] - U^{0,c} \leq 0$; $E[Y|Z=0, D=0] - E[Y|Z=1, D=1] \leq 0$; and $E[Y|Z=0, D=0] - E[Y|Z=0, D=1] \leq 0$. Hence, $LNATE_c \geq \max\{E[Y|Z=1, D=0] - U^{0,c}, E[Y|Z=1, D=0] - E[Y|Z=0, D=$

$1], 0\} = \max\{0, \overline{L}^{c}\}$, with $\overline{L}^{c}$ as defined in the proof of Prop. 3. Using the bounds previously derived for the components of $LNATE_c$, we have that the upper bound is $\widetilde{U}^c \equiv E[Y|Z = 1, D = 1] - E[Y|Z = 0, D = 0]$.

Bounds for $LMATE_c$: $LMATE_c = E[Y^z(1)|c] - E[Y(1, D(0))|c]$. Ass. 4.1 directly implies $LMATE_c \geq 0$. Using the bounds previously obtained for the components of $LMATE_c$ we obtain three additional potential lower bounds: $L^{1,c} - E[Y|Z = 1, D = 1]$, $E[Y|Z = 1, D = 0] - E[Y|Z = 1, D = 1]$ and $E[Y|Z = 0, D = 0] - E[Y|Z = 1, D = 1]$. Each of these three expressions is less than or equal to zero. Using the bounds previously derived for the components of $LMATE_c$ we have the upper bound is $\widetilde{U}^c_m \equiv E[Y|Z = 1, D = 1] - \max\{E[Y|Z = 1, D = 0], E[Y|Z = 0, D = 0]\}$. Thus, $0 \leq LMATE_c \leq \widetilde{U}^c_m$.

We now derive the bounds for $MATE$, starting with the lower bound. As before, we use equations (10) to (13) and the bounds obtained above to derive potential lower bounds for $MATE$. The corresponding four potential lower bounds are:

$$
\begin{aligned}
\Delta_1 &= 0 \\
\Delta_2 &= p_{0|1}L^{0,nt} + p_{1|0}E\left[Y|Z=0, D=1\right] - \left(p_{1|1} - p_{1|0}\right)\widetilde{U}^c - E\left[Y|Z=0\right] \\
&\quad + \left(p_{1|1} - p_{1|0}\right)\max\{L^{1,c}, E[Y|Z=1, D=0], E[Y|Z=0, D=0]\} \\
\Delta_3 &= E\left[Y|Z=1\right] - p_{1|0}U^{1,at} - p_{0|1}E\left[Y|Z=1, D=0\right] \\
&\quad - \left(p_{1|1} - p_{1|0}\right)E[Y|Z=1, D=1] \\
\Delta_4 &= E\left[Y|Z=1\right] - E\left[Y|Z=0\right] - p_{1|0}U^{at} - p_{0|1}U^{nt} - \left(p_{1|1} - p_{1|0}\right)\widetilde{U}^c.
\end{aligned}
$$

After some algebra we obtain $\Delta_2 = -\pi_c(E[Y|Z = 1, D = 1] - \max\{L^{1,c}, E[Y|Z = 1, D = 0], E[Y|Z = 0, D = 0]\}) - p_{0|1}(E[Y|Z = 0, D = 0] - L^{0,nt})$. By definition, $E[Y|Z = 0, D = 0] \geq L^{0,nt}$ and $E[Y|Z = 1, D = 1] \geq L^{1,c}$. Also, $E[Y|Z = 1, D = 1] \geq E[Y|Z = 1, D = 0]$ and $E[Y|Z = 1, D = 1] \geq E[Y|Z = 0, D = 0]$. Hence, $\Delta_2 \leq 0$. We also have: $\Delta_3 = -p_{1|0}(U^{1,at} - E[Y|Z = 1, D = 1]) \leq 0$; and $\Delta_4 = -p_{1|0}(U^{1,at} - E[Y|Z = 1, D = 1]) - p_{0|1}(E[Y|Z = 0, D = 0] - L^{0,nt}) \leq 0$. Thus, the lower bound for $MATE$ equals $\Delta_1 = 0$.

Now consider the upper bound for $MATE$. Plugging in the bounds derived above for the corresponding non-point identified terms in equations (10) to (13) yields the following potential upper bounds:

$$
\begin{aligned}
\Upsilon_1 &= \left(p_{1|1}-p_{1|0}\right)\widetilde{U}^c_m \\
\Upsilon_2 &= p_{0|1}\min\{E\left[Y|Z=0,D=0\right],E\left[Y|Z=1,D=0\right]\} \\
&\quad +p_{1|0}E\left[Y|Z=0,D=1\right]+\left(p_{1|1}-p_{1|0}\right)E[Y|Z=1,D=1] \\
&\quad -\left(p_{1|1}-p_{1|0}\right)\max\{0,\overline{L}^c\}-E\left[Y|Z=0\right] \\
\Upsilon_3 &= E\left[Y|Z=1\right]-p_{1|0}\max\{E\left[Y|Z=1,D=1\right],E\left[Y|Z=0,D=1\right]\} \\
&\quad -p_{0|1}E\left[Y|Z=1,D=0\right]-\left(p_{1|1}-p_{1|0}\right)\max\{E[Y|Z=1,D=0],E[Y|Z=0,D=0]\} \\
\Upsilon_4 &= E\left[Y|Z=1\right]-E\left[Y|Z=0\right]-p_{1|0}\max\{0,\overline{L}^{at}\}-p_{0|1}\max\{0,\overline{L}^{nt}\} \\
&\quad -\left(p_{1|1}-p_{1|0}\right)\max\{0,\overline{L}^c\}.
\end{aligned}
$$

After some algebra we obtain $\Upsilon_4-\Upsilon_2=\Upsilon_3-\Upsilon_1=p_{1|0}\min\{\overline{L}^{at},0\}\leq 0$. Υ_3 can be greater or less than Υ_4 depending on the data. As before, it is enough to show one case in which $\Upsilon_4-\Upsilon_3$ is greater than zero and one in which it is less than zero. After some algebra we can write $\Upsilon_4-\Upsilon_3=p_{0|1}E\left[Y|Z=1,D=0\right]+\pi_c\max\{E[Y|Z=1,D=0],E[Y|Z=0,D=0]\}-p_{0|0}E[Y|Z=0,D=0]-p_{0|1}\max\{0,\overline{L}^{nt}\}-\pi_c\max\{0,\overline{L}^c\}$. Let $\overline{L}^{nt}=E[Y|Z=1,D=0]-E[Y|Z=0,D=0]\leq 0$. Then, $\Upsilon_4-\Upsilon_3=p_{0|1}(E\left[Y|Z=1,D=0\right]-E[Y|Z=0,D=0])-\pi_c\max\{0,\overline{L}^c\}\leq 0$. Now let $\overline{L}^{nt}=E[Y|Z=1,D=0]-E[Y|Z=0,D=0]\geq 0$. Then, $\Upsilon_4-\Upsilon_3=\pi_c(\overline{L}^{nt}-\max\{0,\overline{L}^c\})$, which is greater or equal to zero if $\overline{L}^c\leq 0$.[25] Thus, the upper bound for $MATE$ equals $\min\{\widetilde{U}^1,\widetilde{U}^2\}$, where $\widetilde{U}^1\equiv\Upsilon_3$ and $\widetilde{U}^2\equiv\Upsilon_4$.

Finally, the bounds for $LATE$ follow directly from the bounds for $MATE$ and the result in Proposition 1. Q.D.E.

[25] In fact, it can be shown that $\Upsilon_4-\Upsilon_3=\pi_c(\overline{L}^{nt}-\max\{0,\overline{L}^c\})\geq 0$ regardless of the value of $\overline{L}^c$ as long as $\overline{L}^{nt}\geq 0$.

Table 2. Point Estimates of Interest

Parameters	***Estimate***		***Standard Error***	
Average Treatment Effects				
ATE of program status on employment	0.041		*(0.011)*	
ATE of program status on earnings	18.11		*(4.759)*	
ATE of program status on degree attainment	0.209		*(0.011)*	
LATE of degree attainment on employment	0.194		*(0.055)*	
LATE of degree attainment on earnings	86.63		*(22.769)*	
Strata Proportions				
π_{nt}	0.34		*(0.007)*	
π_{at}	0.45		*(0.009)*	
π_{c}	0.21		*(0.011)*	
	Employment		**Weekly Earnings**	
Conditional Means	*Estimate*	*Std. Error*	*Estimate*	*Std. Error*
E[Y|Z=0]	0.61	*(0.009)*	170.85	*(3.703)*
E[Y|Z=1]	0.65	*(0.007)*	188.96	*(2.933)*
Testable Implications				
E[Y|Z=0, D=1]-E[Y|Z=0, D=0]	0.09	*(0.018)*	48.87	*(7.365)*
E[Y|Z=1, D=1]-E[Y|Z=1, D=0]	0.15	*(0.014)*	70.50	*(5.894)*
E[Y|Z=1, D=1]-E[Y|Z=0, D=0]	0.13	*(0.015)*	64.23	*(5.708)*

Notes: The labor market outcomes (*Y*) are either weekly earnings or employment status in quarter 12 after randomization. The treatment (*D*) is the attainment of a high school, GED, or vocational degree. The instrumental variable (*Z*) is an indicator for whether the individual was randomly assigned to participate in JC ("program status"). Sample size is 8,020 individuals: 2,975 with *Z* =0 and 5,045 with *Z* =1. Standard errors are based on 5,000 bootstrap replications.

Table 3. Estimated Bounds on the Effect of Degree Attainment on Labor Market Outcomes Using Program Status as an Invalid Instrument

	Proposition 2		Proposition 3		Proposition 4	
Assumptions:	1, 2, 3, and 4		1, 2, 3, and 5		1, 2, 3, 4 and 5	
Parameters	LB	UB	LB	UB	LB	UB
			Employment Outcome			
1. MATE	0.000	0.041	-0.094	0.032	0.000	0.021
	(0.000)	*(0.011)*	*(0.006)*	*(0.003)*	*(0.000)*	*(0.007)*
2. LATE	0.000	0.194	-0.448	0.151	0.000	0.100
	(0.000)	*(0.055)*	*(0.013)*	*(0.014)*	*(0.000)*	*(0.035)*
			Weekly Earnings Outcome			
3. MATE	0.00	18.11	-35.52	14.74	0.00	11.26
	(0.00)	*(4.76)*	*(3.51)*	*(1.46)*	*(0.00)*	*(2.68)*
4. LATE	0.00	86.63	-169.89	70.50	0.00	53.87
	(0.00)	*(22.77)*	*(9.38)*	*(5.89)*	*(0.00)*	*(12.48)*

Notes: MATE, the mechanism average treatment effect, is the part of the average treatment effect of the program status on labor market outcomes that works through the attainment of a degree. LATE is the local average treatment effect of degree attainment on labor market outcomes for those individuals assigned to the program and whose degree attainment is affected by their program status. The labor market outcomes (Y) are either employment status or weekly earnings in quarter 12 after randomization. The treatment (D) is the attainment of a high school, GED, or vocational degree. The instrumental variable (Z) is an indicator for whether the individual was randomly assigned to participate in JC ("program status"). Sample size is 8,020 individuals: 2,975 with Z=0 and 5,045 with Z=1. In parenthesis are standard errors computed as described in footnote 12 in the text.