Bounds on Population Average Treatment Effects with an Instrumental Variable*

Xuan Chen† Carlos A. Flores‡ Alfonso Flores-Lagunes§

February, 2014

Abstract

We derive nonparametric sharp bounds on population average treatment effects with an instrumental variable. The main assumptions employed are monotonicity in the treatment of the average outcomes of specified subpopulations, and mean dominance assumptions across the potential outcomes of these subpopulations. Importantly, the direction of the mean dominance assumptions can be informed from data and some of our bounds do not require an outcome with a bounded support. We illustrate the informativeness of our bounds by assessing the effect of the Job Corps program. Our empirical results suggest that enrolling into the program increases weekly earnings and employment by at least $24.61 and 4.3 percentage points, respectively, and decreases yearly dependence on public welfare benefits by at least $84.29.

Key words and phrases: Sharp Bounds; Average treatment effects; Instrumental variables

JEL classification: C13, C21, J30

*We are grateful for comments from seminar participants at The Seventh New York Econometrics Camp at Syracuse University, and Midwest Econometrics Goup Meetings at University of Kentucky.

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

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

§aflores@binghamton.edu. Department of Economics, State University of New York at Binghamton, and IZA.

1 Introduction

Instrumental variable (IV) approaches have been widely used in the literature of program evaluation due to its high internal validity. An influential framework for studying causality within the IV framework was developed by Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996) (hereafter IA and AIR, respectively). They show that allowing heterogeneous effects, IV estimators point identify the local average treatment effect ($LATE$) for compliers, those whose treatment status is affected by the instrument. A common criticism of their approach is the focus on the effect for a subpopulation (e.g., Heckman, 1996; Robins and Greenland, 1996; Deaton, 2010; Heckman and Urzua, 2010), and the instrument-specific interpretation of the $LATE$ stimulates recently growing literature on IV approaches in pursuit of external validity. Point identification of the population treatment effects usually requires an instrument to be strong enough to drive the probability of being treated from zero to one (e.g., Heckman, 2010), which is hard to satisfy in practice. An alternative way relies on stable IV estimates revealed empirically, which inspire the use of multiply instruments for the same causal relationship (e.g., Angrist and Fernandez-Val, 2010). In contrast with point identification, Manski (1990) pioneered partial identification of the population average treatment effect (ATE) under the mean independence assumption of the instrument.

Our paper derives nonparametric sharp bounds for the population ATE by extending the work of IA and AIR. There has been a growing literature on partial identification of the ATE with IV methods since Manski (1990). One strand of this literature endeavors to improve Manski's bounds by assuming different versions of monotonicity of the outcome. Manski and Pepper (2000) introduce the monotonicity of the treatment response (MTR) assumption and the monotonicity of the treatment selection (MTS) assumption. Combining with the mean independence assumption, Chiburis (2010a) derives the bounds under both MTR and MTS assumptions without specifying the direction of the monotonicity a priori. Another strand of the literature imposes structural models on the treatment or the outcome. Under the statistical independence assumption of the instrument, Heckman and Vytlacil (2000) impose a threshold crossing model with a separable error on the treatment. Focusing on a binary outcome, Shaikh and Vytlacil (2011) impose threshold crossing models on both the treatment and the outcome, while Chiburis (2010b) considers a threshold crossing model on the outcome. Instead of assuming the threshold crossing model with separable errors, Chesher (2010) imposes a non-separable structural model on the outcome and assumes the structural function is weakly increasing in the non-separable error.

Comparison of identification power among these assumptions are also discussed in the existing literature on partial identification with IV methods. The assumptions employed by IA and AIR are also cited despite their purpose of point identification. First, the monotonicity assumption on the treatment (e.g., IA; AIR; Balke and Pearl, 1997; Huber and Mellace, 2010) and the structural model assumptions on the treatment (e.g., Heckman and Vytlacil, 2000) do not improve Manski's bounds derived under the mean independence assumption. Second, monotonicity assumptions on the outcome (e.g., Manski and Pepper, 2000) and the structural model assumptions on the outcome do improve Manski's bounds (e.g., Bhattacharya, Shaikh and Vytlacil, 2008, hereafter BSV; Chiburis, 2010a; Chiburis, 2010b; Chesher 2010; Shaikh and Vytlacil, 2011). Third, partial identification with IV methods usually requires bounded support of the outcome. This might also be the reason why quite a few papers focus on binary outcomes (e.g., Balke and Pearl, 1997; BSV; Hahn, 2010; Chiburis, 2010b; Shaikh and Vytlacil, 2011).

It's worth noting that for a binary dependent variable, the monotonicity assumptions and the structural model assumptions are equivalent. Vytlacil (2002) shows the equivalence between the monotonicity assumption and the threshold crossing model on the treatment. Machado et al. (2011) notice the equivalence between the MTR assumption and the threshold crossing model on the outcome. In the absence of covariates, Chiburis (2010b) observe the equivalence between the threshold crossing model with a separable error and the non-separable structural function being weakly increasing in the non-separable error. BSV show that in the absence of covariates, the bounds for a binary outcome under MTR and the mean independence assumptions are equal to the ones derived from the threshold crossing models on both the treatment and the outcome. Afterwards, Chiburis (2010b) notice that his bounds under the threshold crossing model only on the outcome are equal to the ones under MTR and the mean independence assumptions.

This paper improves Manski's nonparametric bounds by extending the work of IA and AIR. We consider the setup consisting of a binary instrument and a binary treatment, which is quite common in the existing literature on partial identification of the ATE with IV methods. We add to the literature by considering two different sets of assumptions that can be useful in practice. The first is monotonicity in the treatment of the average outcomes of subpopulations (strata) defined by the joint potential values of the treatment status under each value of the instrument. Just as BSV and Shaikh and Vytlacil (2011), we do not require prior knowledge about the direction of the monotonicity. However, in contrast to the existing literature (e.g., Manski and Pepper, 2000; BSV; Shaikh and Vytlacil, 2011), we impose monotonicity on the average outcomes of the strata rather than on the individuals' outcomes. This makes the assumption more plausible in practice by allowing some individuals to experience a treatment effect that has the opposite sign to the ATE. The second set of assumptions involves mean dominance assumptions across the potential outcomes of different strata, which have been shown to have significant identifying power in other settings (e.g., Zhang, Rubin and Mealli, 2008; Flores and Flores-Lagunes, 2010). The direction of the mean dominance assumptions can be informed by comparing average baseline characteristics across strata that are likely to be highly correlated with the outcome. Importantly, some of our bounds do not require a bounded outcome.

A recent paper by Huber and Mellace (2010) also derives bounds on the ATE within the IV framework. The main difference between our work and theirs is that we consider the monotonicity assumption on the average outcomes of the strata, which results in narrower bounds and can be justified by economic theory in many applications. Also, we avoid imposing the direction of the monotonicity a priori, while its direction can be inferred from data. In addition, the mean dominance assumptions we consider not only differ from theirs, but the direction of our mean dominance can be informed by comparing the average baseline characteristics across strata, which are estimated by solving an overidentified nonparametric GMM problem.

To illustrate the identification power of our bounds, we analyze the effect of enrolling into the Job Corps program, which is the largest federally-funded job training program for disadvantaged youth in the United States. Using randomization into the program as an instrument, the narrowest bounds on the ATEs derived by our assumptions are [24.61, 201.04] for weekly earnings, [.042, .163] for employment, and $[-142.76, -84.29]$ for public benefits. These bounds are significantly narrower than the IV bounds proposed by Manski (1990), Heckman and Vytlacil (2000) and Kitagawa (2009), when applied to our setting, and the ones by Huber and Mellace (2010). The width of our bounds is also smaller than that under the IV and MTR assumptions of Manski and Pepper(2000), especially for public benefits. Our bounds on employment are also narrower than the ones proposed by Balke and Pearl (1997), BSV, Chesher (2010), Chiburis (2010b) and Shaikh and Vytlacil (2011) for the case of a binary outcome. Our lower bounds for weekly earnings and employment are 10 percent higher than their respective intention-to-treat (ITT) effects (22.19 and .038), while the upper bound for public benefits is equal to its ITT effect.[1] Meanwhile, the $LATE$s for compliers on the three outcomes fall within our narrowest bounds. In sum, our empirical results suggest that enrolling into the Job Corps program increases weekly earnings by at least \$24.61 and employment by at least 4.3 percentage points, and decreases yearly dependence on public welfare benefits by at least \$84.29.

2 Framework

2.1 Setup and Benchmark Bounds

Consider a random sample of size n from a population. Let $D_i \in \{0, 1\}$ indicate whether unit i is treated ($D_i = 1$) or not ($D_i = 0$), and let $Z_i \in \{0, 1\}$ be an instrument for treatment. Let $D_i(1)$ and $D_i(0)$ denote the treatment individual i would receive if $Z_i = 1$ or $Z_i = 0$, respectively. Our outcome of interest is Y. Denote by $Y_i(1)$ and $Y_i(0)$ the potential outcomes as a function of D, i.e., the outcomes individual i would experience if she received the treatment or not, respectively. Finally, let $Y_i(z, d)$ be the potential outcome as a function of the instrument and the treatment.

[1] This is because our bounds derive different numerical results dependent on the sign of the $LATE$ for compliers.

Our parameter of interest is the population average treatment effect, $ATE = E[Y_i(1) - Y_i(0)]$. For each unit, we observe $\{Z_i, D_i(Z_i), Y_i(Z_i, D_i(Z_i))\}$. We omit the subscript i unless necessary for clarity. This setting has received considerable attention in the literature (e.g., AIR, BSV).

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

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

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

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

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

Assumptions 1 through 3 are standard assumptions in the IV literature (e.g., IA, AIR, Huber and Mellace, 2010; Blanco, Flores, and Flores-Lagunes, 2012). Assumption 4 rules out the existence of defiers (compliers) when the monotonicity is non-decreasing (non-increasing). The direction of the monotonicity can be inferred from the data given the independence of Z. Following BSV, we order Z so that $E[D|Z=1] \geq E[D|Z=0]$ to simplify notation.

To partially identify the ATE, we write it as a function of the average effects for the existing strata. Let $LATE_k = E[Y(1) - Y(0)|k]$ and π_k denote, respectively, the local average treatment effect and the stratum proportion in the population, for stratum k, with $k = at$, nt, c. Let $\overline{Y}^{zd} = E[Y|Z=z, D=d]$ and $p_{d|z} = \Pr(D=d|Z=z)$. Under Assumptions 1 through 4, the following quantities are point identified: $\pi_{at} = p_{1|0}$, $\pi_{nt} = p_{0|1}$, $\pi_c = p_{1|1} - p_{1|0}$, $E[Y(1)|at] = \overline{Y}^{01}$, $E[Y(0)|nt] = \overline{Y}^{10}$ and $LATE_c = (E[Y|Z=1] - E[Y|Z=0])/(p_{1|1} - p_{1|0})$. As shown in IA and AIR, $LATE_c$ is point identified for compliers whose treatment status is affected by the instrument, and equals the conventional IV estimand in the absence of covariates. Decomposing the population ATE as a weighted average of the $LATE$s for always-takers, never-takers, and compliers, we have:

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

where $E[Y|Z=z] = E[E[Y|Z=z, D=d]|Z=z]$ is used in the second equality. By equation (2), since $Y(1)$ for never-takers and $Y(0)$ for always-takers are never observed in the data, additional assumptions are needed to bound ATE. The most basic assumption considered in the previous literature is the bounded support of the outcome.

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

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

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

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

The bounds in Proposition 1, which we present for reference, coincide with the IV bounds in Manski (1990), Heckman and Vytlacil (2000) and Kitagawa (2009), when applied to our setting, and with those in Huber and Mellace (2010). When the outcome is binary, these bounds also coincide with those in Balke and Pearl (1997).

2.2 Bounds under Monotonicity

Now let us introduce the monotonicity assumption we employ to improve the identification power of the bounds in Proposition 1.

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

Assumption 6 requires that the $LATE$s of the three existing strata are all either non-negative or non-positive. This assumption is similar to that in BSV, with the important distinction that we impose it on the $LATE$s rather than on the individual effects, which makes it more plausible in practice by allowing some individuals to have a treatment effect of the sign different from that of the ATE. Since we ordered Z so that $E[D|Z = 1] \geq E[D|Z = 0]$, the direction of the monotonicity is identified from the sign of the IV estimand ($LATE_c$) under the current assumptions. The following proposition presents sharp bounds on the ATE under the additional Assumption 6.

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

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

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

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

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

2.3 Bounds under Mean Dominance

In practice, some strata are likely to have more favorable characteristics and thus better mean potential outcomes than others. The three alternative assumptions below formalize the notion that under the same treatment status, never-takers tend to have the best average potential outcome among the three strata, while always-takers tend to have the worst one.

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

[2] For a discussion of the trade-off between the MTR assumption and assuming monotonicity of the treatment in the instrument, see BSV.

The direction of these assumptions can be inverted depending on the application in question. The always-takers and never-takers are likely to be the most "extreme" groups in many applications, so Assumption 7a may be viewed as the weakest of the three. Assumption 7b compares the mean $Y(0)$ and $Y(1)$ of the always-takers and never-takers, respectively, to those of a weighted average of the other two strata, while Assumption 7c compares them to those of the compliers. Although none of these assumptions is directly testable, it is possible to obtain indirect evidence about their plausibility by comparing relevant average pre-treatment characteristics of the different strata (e.g., Flores and Flores-Lagunes, 2010; Frumento et al., 2012). For Assumption 7c, the direction may also be inferred by comparing point identified quantities, $E[Y(1)|at]$ to $E[Y(1)|c]$ and $E[Y(0)|nt]$ to $E[Y(0)|c]$, if these inequalities also hold under the alternative treatment status.

We present bounds under Assumptions 1 through 5 and each of the three versions of Assumption 7. In each case, the lower bound is higher than that in Proposition 1.

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

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

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

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

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

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

Assumptions 7a through 7c have testable implications when combined with Assumption 6, if $LATE_c < 0$. The following inequalities are expected to hold: $\overline{Y}^{01} \leq \overline{Y}^{10}$ (7a); $\overline{Y}^{01} \leq \overline{Y}^{00}$ and $\overline{Y}^{11} \leq \overline{Y}^{10}$ (7b); and, $\overline{Y}^{01} \leq E[Y(0)|c]$ and $E[Y(1)|c] \leq \overline{Y}^{10}$ (7c). If some (or all) of these inequalities are not rejected in applications, then their corresponding assumptions are expected to hold. The following three propositions provide bounds when Assumptions 6 and 7 are combined.

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

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

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

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

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

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

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

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

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

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

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

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

Note that, if $LATE_c<0$, the bounds in Propositions 4 through 6 do not require boundedness of the outcome, because Assumption 6 improves upon the upper bound in Proposition 1, while Assumption 7 improves upon the lower bound. In contrast, if $LATE_c>0$, Assumptions 6 and 7 each improve only upon the lower bound in Proposition 1. The bounds in Propositions 4 through 6 are narrower compared with the bounds in Proposition 2 and the corresponding bounds in Proposition 3. This is because under the combined assumptions, the monotonicity assumption improves upon further either the lower or upper bound in Proposition 3, depending on the sign of $LATE_c$, while the mean dominance assumptions further improve upon the lower bound in Proposition 2.

Proposition 5 overlaps with the bounds recently derived by Chiburis (2010a) under the MTR assumption without specifying a priori direction and the decreasing MTS assumption, as well as the mean independence assumption of the instrument. This is because Assumption 7b coincides with the decreasing MTS assumptions imposed on the counterfactual average outcomes for always-takers and never-takers (i.e., $E[Y(0)|at]$ and $E[Y(1)|nt]$). The form of Chiburis' bounds, however, cannot simplify to Proposition 6, in that his monotonicity assumptions also involve the counterfactual average outcome for the mixture of never-takers and compliers and that for the mixture of always-takers and compliers (i.e., $E[Y(1)|Z=0,D=0]$ and $E[Y(0)|Z=1,D=1]$), which are not involved in our setting.

It is important to note that the bounds in Proposition 6 are also the sharp bounds for ATE if we replace Assumption 7c with the assumption, $E[Y(d)|at] \leq E[Y(d)|c] \leq E[Y(d)|nt]$ for $d=0,1$. However, since $E[Y(d)|c]$ may suffer from the potential issue of a weak IV (i.e., $p_{1|1}-p_{1|0}$ is close to zero), and thus be more difficult to estimate than $E[Y|Z=d,D=d]$. Consequently, the estimated bounds in Proposition 5 may produce narrower confidence intervals than those in Proposition 6 if $p_{1|1}-p_{1|0}$ is close to zero.

2.4 Estimation and Inference

Some of our bounds involve minimum (min) or maximum (max) operators, which create complications for estimation and inference. First, because of the concavity (convexity) of the min (max) function, sample analog estimators of the bounds can be severely biased in small samples. Second, closed-form characterization of the asymptotic distribution of estimators for parameters involving min or max functions are very difficult to derive and, thus, usually unavailable. Furthermore, Hirano and Porter (2012) show that there exist no locally asymptotically unbiased estimators and no regular estimators for parameters that are nonsmooth functionals of the underlying data distribution, such as those involving min or max operators. These issues have generated a growing literature on inference methods for partially identified models of this type (see Tamer, 2010, and the references therein).

We employ the methodology proposed by Chernozhukov, Lee and Rosen (2011) (hereafter CLR) to obtain confidence regions for the true parameter value, as well as half-median unbiased estimators for the lower and upper bounds. The half-median-unbiasedness property means that the upper (lower) bound estimator exceeds (falls below) the true value of the upper (lower) bound with probability at least one half asymptotically. This is an important property because achieving local asymptotic unbiasedness is not possible, implying that "bias correction procedures cannot completely eliminate local bias, and reducing bias too much will eventually cause the variance of the procedure to diverge" (Hirano and Porter, 2012). For details on the procedure of applying CLR's method to our bounds see the Appendix. For the bounds without min or max operators, we use sample analog estimators and construct the confidence regions for the true parameter value proposed by Imbens and Manski (2004).

3 Bounds on Population Average Treatment Effects of Job Corps

3.1 Data

Job Corps (JC) is the largest and most comprehensive education and job training program in the United States. It offers to economically disadvantaged youth aged 16 to 24 years old academic education, vocational training, residential living, health care and health education, counseling and job placement assistance. According to the U.S. Department of Labor (2005), a typical JC student lives at a local JC center for eight months and receives about 1100 hours of academic and vocational instruction, which is equivalent to approximately one year in high school.

We employ data from the National Job Corps Study (NJCS), a randomized experiment undertaken in the mid-to-late nineties. The study examined the impacts of JC on labor market outcomes, welfare dependence and several other outcomes to help assess whether the program achieved its goals of helping students become more responsible and productive citizens. Eligible applicants were randomly assigned to treatment and control groups. Individuals in the control group were embargoed from the program for a period of three years. The research sample was interviewed at random assignment and at 12, 30, and 48 months after random assignment. Taking advantage of randomization, most of previous works on JC study ITT effects or $LATE$s for compliers (e.g., Burghardt et al., 2001; Schochet, Burghardt, and Glazerman, 2001; Schochet, Burghardt, and McConnell, 2008; Lee, 2009; Blanco, Flores, and Flores-Lagunes, 2012). The noncompliance behavior, however, tends to dilute the impacts of JC. In our sample, 73% of individuals of the treatment group actually enrolled in JC, while 4% of individuals of the control group also enrolled. Even adjusting to noncompliance by examining $LATE_c$, that effect is representative for a subpopulation accounting for 69% in the population of our interest.

We make inference about the population ATE, which is of great public interest, using data on individuals who responded to the 48-month interview. The outcome variables we consider are weekly earnings and employment at Week 208 and public assistance benefits received during the fourth year after randomization[3]. Given the objectives and services provided by JC (e.g., academic and vocational training, job search assistance), it tends to have positive effects on participants' labor market outcomes, though the direction of these effects may be inverted in the short run. On one hand, individuals who enroll in JC could be less likely to be employed while undergoing training, which is usually referred to as the "lock-in" effect (van Ours, 2004). On the other hand, some participants may raise their reservation wages after training and

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

choose to be unemployed by rejecting some job offers. Both potential threats make participants possibly experience lower earnings than those who do not enroll in JC in the short run. Given a long enough period of time, however, trained individuals are no longer "locked-in" away from employment and those raising reservation wages find jobs, and both of them are more likely to have higher earnings after training. Consistent with this view, Schochet, Burghardt, and Glazerman (2001) and Lee (2009) find negative effects of JC on weekly earnings and employment in the short run, but positive effects in the long run.

Participation in JC may also affect welfare dependence differently in the short and long runs. In the short run, participants may experience a reduction in welfare receipt while they enroll in JC, because the program provides shelter (except to nonresidential students), food, and a stipend. In the long run, after they leave JC, participants may receive less public income support because of higher earnings. In contrast with this expectation, Schochet, Burghardt, and Glazerman (2001) report that the reductions in benefit receipt persisted throughout 4 years after randomization. Therefore, we focus our analysis on the latest measures available of labor market outcomes and welfare dependence in the NJCS.

The treatment variable indicates whether or not the individual ever enrolled in JC during the 208 weeks (i.e., four years) after random assignment. The random assignment indicator serves as an instrument for JC enrollment. Two samples are obtained by dropping individuals with missing relevant variables from the survey.[4] The sample for weekly earnings and employment involves 10,520 individuals (4,187 and 6,333 in the control and treatment groups, respectively), while for public benefits 10,976 individuals (4,387 and 6,589 in the control and treatment groups, respectively). Finally, due to both design and programmatic reasons, some subpopulations were randomized in the NJCS with different (but known) probabilities (Schochet, Burghardt, and Glazerman, 2001). Hence, we employ design weights throughout our analysis.[5]

Table 1 reports the average baseline characteristics of both samples by treatment assignment status along with the percentage of missing values for each of those variables. The pre-treatment variables include demographic characteristics, education and background variables, employment, earnings and public benefits dependency at baseline, as well as labor market outcomes in the year prior to randomization. As one would expect, the average pre-treatment characteristics of the treatment and control groups are similar in both our samples due to randomization, with the difference in means being statistically different from zero at the five percent level for only one variable (personal income: 3,000-6,000). Thus, both our samples maintain the balance of baseline variables between the control and treatment groups.

[4] We derive two samples because individuals with missing labor market outcomes and with missing public benefits are different.

[5] Specifically, the weight we employ addresses sample design, 48-month interview design, and 48-month interview non-response.

3.2 Assessment of Assumptions

Table 2 shows some relevant point identified averages for both samples. The noncompliance behavior is similar between the two samples. As we have already mentioned, 73% of individuals of the treatment group actually enrolled in JC, while 4% of individuals of the control group also enrolled during the 208 weeks after randomization. The ITT effects on weekly earnings, employment and public benefits are 22.19 , .038 and -84.29, respectively. These effects are all statistically significant, with their signs as expected. The $LATE_c$ estimates for compliers on earnings, employment and public benefits are 32.29, .055 and -122.28, respectively, 45 percent higher than their corresponding ITT estimates. By Assumption 6, the sign of $LATE_c$ identifies the sign of the $LATE$s for the other two strata. Thus, our estimates of $LATE_c$ indicate positive population average treatment effects on weekly earnings and employment and a negative population effect on public benefits.

The middle part of Table 2 shows the proportion of each stratum. In both samples, the proportion of compliers is the largest, .69, followed by never-takers, .27, and always-takers, .04. And by Assumption 6, there are no defiers in our samples. The end part of Table 2 reports the point identified averages cited in Assumptions 6 and 7.[6] These estimates are all statistically significant and follow a certain pattern in both samples: under the treated status, the average outcome for always-takers is the smallest, followed by the average for the mixture of always-takers and compliers, and the average for compliers, while under the untreated status, the average outcome for compliers is the smallest, followed by the average for the mixture of never-takers and compliers, and the average for never-takers.

As mentioned previously, differences across these point identified averages may provide a preliminary hint for Assumption 7. To begin, we may infer the direction of Assumption 7c by comparing the identified averages of always-takers and never-takers to those of compliers under the same treatment status. The hypotheses that $E[Y(1)|at] \leq E[Y(1)|c]$ and $E[Y(0)|c] \leq E[Y(0)|nt]$ are not rejected for all of the three outcomes. Thus, if the same relationship with compliers also hold under the alternative treatment status, we may expect Assumption 7c to hold. Furthermore, since the ITT effect on public benefits is negative, testable implications are available when Assumption 7 is combined with Assumption 6, as discussed in Section 2.3. These testable implications are not rejected in our application.

More importantly, indirect evidence of Assumption 7 are obtained by estimating the average baseline characteristics across strata from an overidentified nonparametric GMM problem. For details on the GMM procedure see Appendix. Tables 3 and 4 show these estimates and their differences across strata for our samples. The average characteristics across strata are similar

[6] As in Lee (2009), we use a *transformed measure* to estimate the sample averages of weekly earnings and public benefits to minimize the effect of outliers. Specifically, the entire observed outcome distribution (for either weekly earnings or public benefits) is split into 20 percentile groups ($5^{th}, 10^{th}, \ldots, 95^{th}$), and then the mean outcome within each of the 20 groups is assigned to each individuals.

between the two samples. Among the three strata, never-takers are more likely to be female, older, married, to have children, higher level of education, personal income above $9,000 (less likely to have personal income below $3,000), higher weekly earnings at baseline, and to have better labor market outcomes the year before randomization. By contrast, always-takers tend to be male, younger, to have lower level of education at baseline, and to have lower earnings in the previous year. Furthermore, the statistical significance of the difference between always-takers and never-takers indirectly supports Assumption 7a, while the statistical significance of the differences compared to compliers (i.e., columns 4 and 5) tends to support Assumption 7c. Meanwhile, when the differences across the three strata are all statistically significant, Assumption 7b are more likely to hold. Note that these differences become more convincing evidence of Assumption 7 on the labor market outcomes under the untreated arm than those under the treated arm. However, the differences across the strata in the public benefit dependency at the baseline are not statistically significant.[7] We conclude from these results that the data do not provide indirect evidence against Assumption 7, and that the point estimates of the differences suggest that the assumption is plausible.

3.3 Empirical Results

Table 5 shows our bounds on the ATEs on the labor market outcomes and the public dependency under Proposition 1 through Proposition 6. Under each pair of the estimated bounds, we report a 95% level confidence interval for the true parameter. Since the bounds for weekly earnings and employment in Propositions 4 through 6 involve max or min operators, we report the half-median unbiased estimators and the corresponding confidence intervals proposed by CLR. The bounds without max or min operators are estimated with sample analogs and their confidence intervals are obtained by the method of Imbens and Manski (2004).

We begin with the ATE on weekly earnings in the first two columns. Proposition 1 provides the bounds in the AIR setting under the bounded-outcome assumption (A.5). The estimated bounds are rather wide and fail to identify the sign of the ATE. Note that these bounds are also the IV bounds proposed by Manski (1990), Heckman and Vytlacil (2000) and Kitagawa (2009), when applied to our setting, and the ones by Huber and Mellace (2010). The ATE in Proposition 2 under the monotonicity assumption (A.6) is bounded between [22.19, 201.02], obtained by identifying positive $LATE$s for always-takers and never-takers. Note that they are also the ones under the IV and MTR assumptions proposed by Manski and Pepper (2000). The mean dominance assumptions (A.7) improve upon the lower bound in Proposition 1, and negative effects are ruled out in Propositions 3b and 3c in the absence of assuming the sign of $LATE$s. When we impose Assumptions 1 through 7 together, all of our bounds and their corresponding confidence intervals lie in the positive region. In particular, the bounds on the

[7] Unfortunately, information about the amount of public benefits in dollars is unavailable at the baseline.

ATE on the weekly earnings in Proposition 6 are the narrowest, [24.61, 201.04], with the lower bound 10 percent higher than the ITT effect (22.19), while the $LATE_c$ for compliers (32.29) falls within the bounds. It turns out that informing the unobserved terms (i.e., $E[Y(0)|at]$ and $E[Y(1)|nt]$) from the point identified outcomes of compliers provide a sharper lower bound on the ATE on the weekly earnings than that obtained by identifying the sign of $LATE_{at}$ and $LATE_{nt}$ under the monotonicity assumption (A.6).

The next two columns show the bounds on the ATE on the employment, whose value is between 0 and 1. A similar pattern to the bounds for the weekly earnings is also found in the bounds for the employment. Note that the bounds in Proposition 1, [−.015, .163], unable to identify the sign of the ATE, also coincide with those in Balke and Pearl (1997) for the case of a binary outcome. The bounds in Proposition 2 derive a positive ATE, which varies between [.038, .163], and also equal the ones proposed by BSV, Chesher (2010), Chiburis (2010b) and Shaikh and Vytlacil (2011), all of which analyze a binary outcome. Again, Proposition 6 provides the narrowest bounds on the ATE on the employment under Assumptions 5, 6 and 7c, [.042, .163], with the lower bound 10 percent higher than the ITT effect (.038), while the $LATE_c$ for compliers (.055) falls within the bounds in Proposition 6.

The final two columns report the bounds on the ATE on the public benefits. Different from the labor market outcomes, since the ITT effect on the public benefits is negative, imposing only the monotonicity assumption improves upon its upper bound in Proposition 1, while imposing only the mean dominance assumptions improves upon its lower bound. The bounds in Proposition 1 are extremely wide and uninformative. The monotonicity assumption (A.6) has a strong identification power compared with the mean dominance assumptions (A.7) in the case of the public benefits, in that the former identifies the negative sign of the ATE on the public benefits, though the latter greatly sharpen the lower bound in Proposition 1 by at least 70 percent. However, once we consider the two types of assumptions together, the bounded outcome assumption is no longer necessary and the width of the bounds shrink dramatically. Under the combined assumptions, the estimated bounds and their corresponding confidence intervals lie in the negative region. Proposition 6 provides the narrowest bounds on the ATE on public benefits, [−142.76, −84.29], with the upper bound equal to the ITT effect, while the $LATE_c$ on compliers (−122.28) falls within the bounds.

4 Conclusion

This paper derives sharp nonparametric bounds on the population average treatment effects by extending the work of Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996). The favorable bounds are derived by combining two sets of assumptions. The first is monotonicity in the treatment of the average outcomes of strata without specifying a priori direction. The second is mean dominance assumptions on average potential outcomes across strata. Importantly, some of our bounds do not require a bounded support. And empirically, the direction of the mean dominance assumption can be inferred from data by estimating an overidentified nonparametric GMM problem. The application to the Job Corps program illustrates the informativeness of our bounds. Our empirical results suggest that enrolling into the program increases weekly earnings by at least \$24.61 and employment by at least 4.3 percentage points at Week 208 after randomization, and decreases the dependence on public welfare benefits by at least \$84.29 during the fourth year after randomization.

We close by pointing out that a similar analytical strategy to the one followed here can be used to bound the ATE when the instrument does not satisfy the exclusion restriction, in which case the local average treatment effect for compliers can be bounded as in Flores and Flores-Lagunes (2012). In addition to solving the endogeneity issue, the strategy can also be applied to the identification of direct and indirect effects (e.g., Rubin, 2004; Sjölander, 2009; VanderWeele, 2011)

Table 1: Summary Statistics of Baseline Variables

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

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

Table 2: Point Identified Average Outcomes after Random Assignment

Variables:	Sample for Labor Market Outcomes at Week 208 Enrollment	Earnings	Employment	Sample for Public Benefits in Year 4 Enrollment	Public benefits
Averages for $Z = 1$	.730**	228.78**	.608**	.732**	747.21**
	(.006)	(3.004)	(.006)	(.005)	(23.40)
Averages for $Z = 0$	.043**	206.60**	.570**	.043**	831.50**
	(.003)	(3.552)	(.008)	(.003)	(30.28)
$ITT\ Effects$	.687**	22.19**	.038**	.689**	-84.29**
	(.006)	(4.652)	(.010)	(.006)	(38.27)
$LATE_c$		32.29**	.055**		-122.28**
		(7.007)	(.015)		(56.78)
Proportions of Strata under Assumptions 1 to 4					
π_{nt}	.270**			.268**	
	(.006)			(.006)	
π_c	.687**			.689**	
	(.007)			(.006)	
π_{at}	.043**			.043**	
	(.003)			(.003)	
Other Point Identified Average Outcomes under Assumptions 1 to 4					
$E[Y(1)\|at]$		132.10**	.393**		545.45**
		(14.94)	(.037)		(110.12)
$E[Y(0)\|nt]$		223.79**	.600**		880.67**
		(5.967)	(.012)		(47.98)
$E[Y(1)\|c]$		236.82**	.624**		707.81**
		(4.022)	(.008)		(28.26)
$E[Y(0)\|c]$		204.53**	.569**		830.09**
		(5.655)	(.012)		(49.69)
$E[Y\|Z = 1, D = 1]$		230.63**	.611**		698.35**
		(3.614)	(.007)		(25.87)
$E[Y\|Z = 0, D = 0]$		209.96**	.578**		844.25**
		(3.709)	(.008)		(33.18)

Note: Z denotes whether the individual was randomly assigned to participate ($Z = 1$) or not ($Z = 0$) in the program. D denotes whether the individual was ever enrolled in the program ($D = 1$) or not ($D = 0$) during the 4 years (208 weeks) after randomization. Benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Numbers in parentheses are standard errors. ** denotes estimate is statistically different from 0 at 5% level. Computations use the weights that account for sample and interview design and interview non-response. The standard errors of $LATE$s, proportions of strata and other identified average outcomes are calculated by 5000-repetition bootstrap.

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

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

Note: Benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Numbers in parentheses are standard errors. ** and * denote that estimate is statistically different from 0 at 5% and 10% level, respectively. Computations use the weights that account for sample and interview design and interview non-response. Missing values for each of the baseline variables were imputed with the mean of the variable.

Table 4: Average Baseline Characteristics in the Sample for Public Assistance Benefits

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

Note: Benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Numbers in parentheses are standard errors. ** and * denote that estimate is statistically different from 0 at 5% and 10% level, respectively. Computations use the weights that account for sample and interview design and interview non-response. Missing values for each of the baseline variables were imputed with the mean of the variable.

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

	Earnings		Employment		Public Benefits	
	LB	UB	LB	UB	LB	UB
Bounds under Each Assumption						
Proposition 1	-69.86	201.02	-.150	.163	-632.86	1812.4
Bounded Outcome (A.5)	(-78.34, 210.61)		(-.167, .179)		(-702.21, 1901.6)	
Proposition 2	22.19	201.02	.038	.163	-632.86	-84.29
Monotonicity (A.6)	(14.18, 210.61)		(.021, .179)		(-702.21, -22.13)	
Proposition 3a	-6.507	201.02	-.027	.163	-188.43	1812.4
Mean Dominance (A.7a)	(-16.65, 210.61)		(-.050, .179)		(-265.90, 1901.6)	
Proposition 3b	20.67	201.02	.033	.163	-145.90	1812.4
Mean Dominance (A.7b)	(11.97, 210.61)		(.015, .179)		(-212.69, 1901.6)	
Proposition 3c	22.57	201.02	.037	.163	-142.76	1812.4
Mean Dominance (A.7c)	(13.72, 210.61)		(.019, .179)		(-210.62, 1901.6)	
Bounds under Combined Assumptions						
Proposition 4	20.43	201.02	.034	.163	-188.43	-84.29
(A.5, A.6, A.7a)	(13.01, 210.58)		(.018, .180)		(-265.95, -22.09)	
Proposition 5	22.97	201.01	.039	.163	-145.90	-84.29
(A.5, A.6, A.7b)	(14.53, 210.56)		(.020, .180)		(-213.01, -21.83)	
Proposition 6	24.61	201.04	.042	.163	-142.76	-84.29
(A.5, A.6, A.7c)	(16.01, 210.59)		(.023, .180)		(-210.62, -22.13)	

Note: Benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Numbers in parentheses are 95% level confidence intervals for true parameters of interest. The confidence intervals of the bounds under each assumption are calculated by the method of Imbens and Manski (2004). For earnings and employment, the confidence intervals of the bounds under combined assumptions are calculated by the method of Chernozhukov, Lee and Rosen (2011), while the bounds under combined assumptions are estimated by the half-median unbiased estimators proposed by Chernozhukov, Lee and Rosen (2011). For public benefits, the confidence intervals are calculated by the method of Imbens and Manski (2004). Computations use the weights that account for sample and interview design and interview non-response. Standard errors are calculated by 5000-repetition bootstrap.

References

[1] Angrist, J. (1990), "Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records," American Economic Review 80, 313-335.

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

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

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

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

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

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

[8] Chernozhukov, V., Lee, S. and Rosen, A. (2011), "Intersection bounds: estimation and inference," forthcoming, Econometrica.

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

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

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

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

[13] Flores, C. and Flores-Lagunes, A. (2010), "Partial identification of local average treatment effects with an invalid instrument," Working Paper.

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

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

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

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

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

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

[20] Huber, M. and Mellace, G. (2010), "Sharp IV bounds on average treatment effects under endogeneity and noncompliance," Working Paper.

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

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

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

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

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

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

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

[28] Oreopoulos, P. (2006), "Estimate average and local average treatment effects of education when compulsory schooling laws really matter," American Economic Review 96, 152-175.

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

[30] Rosenzweig, M. and Wolpin, K. (1980), "Testing the quantity-quality fertility model: the use of twins as a natural experiment," Econometrica, 48, 227-240.

[31] Rubin, D. (2004), "Direct and indirect causal effects via potential outcomes (with discussion)," Scandinavian Journal of Statistics 31, 161-198.

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

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

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

[35] Sjölander, A. (2009), "Bounds on natural direct effects in the presence of confounded intermediate variables," Statistics in Medicine 28, 558-571.

[36] Tamer, E. (2010), "Partial identification in econometrics," Annual Review of Economics 2, 167-195.

[37] U.S. Department of Labor (2005), "Job Corps fact sheet," http://www.doleta.gov /Programs/factsht/jobcorps.cfm (December 24, 2006).

[38] van Ours, J. (2004), "The locking-in effect of subsidized jobs," Journal of Comparative Economics 32, 37-55.

[39] VanderWeele, T.J. (2011), "Causal mediation analysis with survival data," Epidemiology 22, 582-585.

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

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

A Appendix

A.1 Proof

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

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

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

and

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

$ATE = E[Y(1) - Y(0)]$

$= E[Y(1)|Z = 1] - E[Y(0)|Z = 0]$

$= p_{1|1}E[Y(1)|Z = 1, D = 1] + p_{0|1}E[Y(1)|Z = 1, D = 0] - p_{1|0}E[Y(0)|Z = 0, D = 1] - p_{0|0}E[Y(0)|Z = 0, D = 0]$

$= p_{1|1}E[Y|Z = 1, D = 1] + p_{0|1}E[Y(1)|Z = 1, D = 0] - p_{1|0}E[Y(0)|Z = 0, D = 1] - p_{0|0}E[Y|Z = 0, D = 0]$

$= p_{1|1}\overline{Y}^{11} + p_{0|1}q_1 - p_{1|0}q_0 - p_{0|0}\overline{Y}^{00}$

$= \alpha.$

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

A.2 GMM Moment Function

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

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

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

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

A.3 Empirical Results on the $LATE$s

Appendix Table A1: Bounds for Earnings at Week 208

	$LATE_{nt}$		$LATE_{at}$		ATE	
	LB_{nt}	UB_{nt}	LB_{at}	UB_{at}	LB	UB
Bounds under Each Assumption						
Proposition 1	-223.79	641.87	-733.57	132.10	-69.86	201.02
Bounded Outcome	(-233.61, 651.70)		(-758.51, 157.03)		(-78.34, 210.61)	
Proposition 2	.000	641.87	.000	132.10	22.19	201.02
Monotonicity	(.000, 651.70)		(.000, 157.03)		(14.18, 210.61)	
Proposition 3a	-91.70	641.87	-91.70	132.10	-6.507	201.02
Mean Dominance	(-118.63, 651.70)		(-118.63, 157.03)		(-16.65, 210.61)	
Proposition 3b	6.834	641.87	-77.86	132.10	20.67	201.02
Mean Dominance	(-4.625, 651.70)		(-103.62, 157.03)		(11.97, 210.61)	
Proposition 3c	13.03	641.87	-72.43	132.10	22.57	201.02
Mean Dominance	(1.260, 651.70)		(-99.15, 157.03)		(13.72, 210.61)	
Bounds under Combined Assumptions						
Proposition 4	.000	641.88	.000	132.08	20.43	201.02
	(.000, 651.70)		(.000, 158.98)		(13.01, 210.58)	
Proposition 5	3.032	641.89	.000	132.12	22.97	201.01
	(.000, 651.70)		(.000, 159.07)		(14.53, 210.56)	
Proposition 6	9.119	641.88	.000	132.12	24.61	201.04
	(.000, 651.70)		(.000, 159.07)		(16.01, 210.59)	

Note: Numbers in parentheses are 95% level confidence intervals for true parameters of interest. The confidence intervals of the bounds under each assumption are calculated by the method of Imbens and Manski (2004). The confidence intervals of the bounds under combined assumptions are calculated by the method of Chernozhukov, Lee and Rosen (2011). The bounds under combined assumptions are estimated by the half-median unbiased estimators proposed by Chernozhukov, Lee and Rosen (2011). Computations use the weights that account for sample and interview design and interview non-response. Standard errors are calculated by 5000-repetition bootstrap.

Appendix Table A2: Bounds for Employment at Week 208

	$LATE_{nt}$		$LATE_{at}$		ATE	
	LB_{nt}	UB_{nt}	LB_{at}	UB_{at}	LB	UB
Bounds under Each Assumption						
Proposition 1	-.600	.400	-.607	.393	-.150	.163
Bounded Outcome	(-.620, .421)		(-.668, .454)		(-.167, .179)	
Proposition 2	.000	.400	.000	.393	.038	.163
Monotonicity	(.000, .421)		(.000, .454)		(.021, .179)	
Proposition 3a	-.207	.400	-.207	.393	-.027	.163
Mean Dominance	(-.272, .421)		(-.272, .454)		(-.050, .179)	
Proposition 3b	.011	.400	-.185	.393	.033	.163
Mean Dominance	(-.013, .421)		(-.247, .454)		(.015, .179)	
Proposition 3c	.025	.400	-.176	.393	.037	.163
Mean Dominance	(.000, .421)		(-.240, .454)		(.019, .179)	
Bounds under Combined Assumptions						
Proposition 4	.000	.400	.000	.393	.034	.163
	(.000, .421)		(.000, .458)		(.018, .180)	
Proposition 5	.003	.400	.000	.393	.039	.163
	(.000, .421)		(.000, .458)		(.020, .180)	
Proposition 6	.017	.401	.000	.393	.042	.163
	(.000, .421)		(.000, .458)		(.023, .180)	

Note: Numbers in parentheses are 95% level confidence intervals for true parameters of interest. The confidence intervals of the bounds under each assumption are calculated by the method of Imbens and Manski (2004). The confidence intervals of the bounds under combined assumptions are calculated by the method of Chernozhukov, Lee and Rosen (2011). The bounds under combined assumptions are estimated by the half-median unbiased estimators proposed by Chernozhukov, Lee and Rosen (2011). Computations use the weights that account for sample and interview design and interview non-response. Standard errors are calculated by 5000-repetition bootstrap.

Appendix Table A3: Bounds for Public Assistance Benefits in Year 4

	$LATE_{nt}$		$LATE_{at}$		ATE	
	LB_{nt}	UB_{nt}	LB_{at}	UB_{at}	LB	UB
Bounds under Each Assumption						
Proposition 1	-880.67	6990.6	-7325.8	545.45	-632.86	1812.4
Bounded Outcome	(-959.40,	7069.3)	(-7506.9,	726.53)	(-702.21,	1901.6)
Proposition 2	-880.67	.000	-7325.8	.000	-632.86	-84.29
Monotonicity	(-959.40,	.000)	(-7506.9,	.000)	(-702.21,	-22.13)
Proposition 3a	-335.22	6990.6	-335.22	545.45	-188.43	1812.4
Mean Dominance	(-531.16,	7069.3)	(-531.16,	726.53)	(-265.90,	1901.6)
Proposition 3b	-182.32	6990.6	-298.80	545.45	-145.90	1812.4
Mean Dominance	(-272.08,	7069.3)	(-488.95,	726.53)	(-212.69,	1901.6)
Proposition 3c	-172.85	6990.6	-284.64	545.45	-142.76	1812.4
Mean Dominance	(-264.71,	7069.3)	(-484.84,	726.53)	(-210.62,	1901.6)
Bounds under Combined Assumptions						
Proposition 4	-335.22	.000	-335.22	.000	-188.43	-84.29
	(-531.16,	.000)	(-531.16,	.000)	(-265.95,	-22.09)
Proposition 5	-182.32	.000	-298.80	.000	-145.90	-84.29
	(-272.08,	.000)	(-488.95,	.000)	(-213.01,	-21.83)
Proposition 6	-172.85	.000	-284.64	.000	-142.76	-84.29
	(-264.71,	.000)	(-484.96,	.000)	(-210.62,	-22.13)

Note: Benefits include AFDC/TANF, food stamps, SSI/SSA, and General Assistance. Numbers in parentheses are 95% level confidence intervals for true parameters of interest. The confidence intervals are calculated by the method of Imbens and Manski (2004). Computations use the weights that account for sample and interview design and interview non-response. Standard errors are calculated by 5000-repetition bootstrap.