Bounds on Average and Quantile Treatment Effects of Job Corps Training on Wages*

German Blanco

Department of Economics, State University of New York at Binghamton

gblanco1@binghamton.edu

Carlos A. Flores

Department of Economics, University of Miami

caflores@miami.edu

Alfonso Flores-Lagunes

Department of Economics, State University of New York at Binghamton

and IZA

aflores@binghamton.edu

June 19, 2012

*We thank Xianghong Li, Oscar Mitnik, and participants at Binghamton University's labor group for detailed comments. We also thank participants at the 2011 Institute for Research in Poverty Summer Workshop, the 2011 Agricultural and Applied Economics Association Meetings, the 2011 Midwest Econometrics Group Meeting, the 2012 Society of Labor Economists Meeting, and seminar participants at Syracuse and York Universities for useful comments.

Abstract

We review and extend recent nonparametric partial identification results for average and quantile treatment effects in the presence of sample selection under various assumptions. These methods are applied to assessing the effect of Job Corps (JC), the largest job training program targeting disadvantaged youth in the United States, on wages. Our preferred estimates, which exclude the group of Hispanics, suggest positive effects of JC on wages both at the mean and throughout the wage distribution. For the different demographic groups analyzed, the statistically significant estimated average effects are bounded between 4.6 and 12 percent, while the statistically significant quantile treatment effects are bounded between 2.7 and 14 percent. We also document that the program's effect on wages varies across quantiles and demographic groups.

1 Introduction

Sample selection is a well-known and commonly found problem in applied econometrics that arises when there are factors simultaneously affecting both the outcome and whether or not the outcome is observed. Sample selection arises, for example, when analyzing the effects of a given policy on the performance of firms, as there are common factors affecting both the performance of the firm and the firm's decision to exit or remain in the market; or when evaluating the effects of an intervention on students' test scores if students can self-select into taking the test. Even in a controlled or natural experiment in which the intervention is randomized, outcome comparisons between treatment and control groups yield biased estimates of causal effects if the probability of observing the outcome is affected by the intervention. For instance, Sexton and Hebel (1984) employ data from a controlled experiment to analyze the effect of an antismoking assistance program for pregnant women on birth weight. Sample selection arises in this context if the program has an effect on fetal death rates. An example of a natural experiment where sample selection bias may arise is on the study of the effects of the Vietnam-era draft status on future health, as draft eligible-men may experience higher mortality rates (Hearst et al., 1986; Angrist et al., 2009; Dobkin and Shabani, 2009; Eisenberg and Rowe, 2009).

In this paper, we review and extend recent nonparametric partial identification results for average and quantile treatment effects in the presence of sample selection. We do this in the context of assessing the wage effects of Job Corps, which is the largest job training program targeting disadvantaged youth in the United Sates.

The vast majority of both empirical and methodological econometric literature on the evaluation of labor market programs focuses on estimating their causal effects on total earnings (e.g., Heckman, LaLonde and Smith, 1999; Imbens and Wooldridge, 2009). Evaluating the impact on total earnings, however, leaves open a relevant question about whether or not these programs have a positive effect on the wages of participants through the accumulation of human capital, which is an important goal of active labor market programs. Earnings have two components: price and quantity supplied of labor. By focusing on estimating the impact of program participation on earnings one cannot distinguish how much of the effect is due to human capital improvements. Assessing the labor market effect of program participation on human capital requires focusing on the price component of earnings, i.e., wages. The reason is that wage increases are directly related to the improvement of participants' human capital through the program. Unfortunately, estimation of the program's effect on wages is not straightforward due to sample selection: wages are observed only for those individuals who are employed (Heckman, 1979). As in the previous examples, randomization of program participation does not solve this problem because the individual's decision to become employed is endogenous and occurs after randomization.

Recently, new partial identification results have been introduced which allow the construction of nonparametric bounds for average and quantile treatment effects that account for sample selection. These bounds typically require weaker assumptions than those conventionally employed for point identification of these effects.[1] We review these techniques and extend them by discussing how to use covariates to narrow the bounds for quantile

[1] Many of the methods employed for point identification of treatment effects under sample selection require strong distributional assumptions that may not be satisfied in practice, such as bivariate normality (e.g., Heckman, 1979). One may relax this distributional assumption by relying on exclusion restrictions (Heckman, 1990; Imbens and Angrist, 1994; Abadie et al., 2002), which require variables that determine selection into the sample (employment) but do not affect the outcome (wages). It is well known, however, that in the case of employment and wages it is difficult to find plausible exclusion restrictions (Angrist and Krueger, 1999; Angrist and Krueger, 2001).

treatment effects. Subsequently, we use data from the National Job Corps Study (NJCS), a randomized evaluation of the Job Corps (JC) program, to empirically assess the effect of JC training on wages. We analyze effects both at the mean and at different quantiles of the wage distribution of participants, as well as for different demographic groups. We focus on estimating bounds for the subpopulation of individuals who would be employed regardless of participation in JC, as previously done in Lee (2009) and Zhang et al. (2008), among others. Wages are non-missing under both treatment arms for this group of individuals, thus requiring fewer assumptions to construct bounds on their effect. This is also an important group of participants: it is estimated to be the largest group among eligible JC participants, accounting for close to 60 percent of them.

We start by considering the Horowitz and Manski (2000) bounds, which exploit the randomization in the NJCS and use the empirical support of the outcome. However, these bounds are wide in our application. Subsequently, we proceed to tighten the bounds through the use of two monotonicity assumptions within a principal stratification framework (Frangakis and Rubin, 2002). The first states individual-level weak monotonicity of the effect of the program on employment. This assumption was also employed by Lee (2009) to partially identify average wage effects of JC. The second assumption (not considered by Lee, 2009) is on mean potential outcomes across strata, which are subpopulations defined by the potential values of the employment status variable under both treatment arms. These assumptions result in informative bounds for our parameters.

We contribute to the literature in two ways. First, we review, extend, and apply recent partial identification results to deal with sample selection. In particular, we illustrate a way to analyze treatment effects on different quantiles of the outcome distribution in the presence of sample selection by employing the set of monotonicity assumptions described above.[2] In addition, we propose a method to employ a covariate to narrow trimming bounds for quantile treatment effects. Second, we add to the literature analyzing the JC training program by evaluating its effect on wages with these new methods. With a yearly cost of about $1.5 billion, JC is America's largest job training program. As such, this federally funded program is under constant examination and, given legislation seeking to cut federal spending, the program's operational budget is currently under scrutiny

[2]Other recent work (to be discussed below) that employs bounds on quantile treatment effects under different monotonicity assumptions are Blundell et al. (2007) and Lechner and Melly (2010).

(e.g., USA Today, 2011). Our results suggest that the program is effective in increasing wages. Moreover, they contribute to a policy-relevant question regarding the potential heterogeneity of the wage impacts of JC at different points of the wage distribution, and across different demographic groups. In this way, we add to a growing literature analyzing the effectiveness of active labor market programs across different demographic groups (Heckman and Smith, 1999; Abadie, Angrist, and Imbens, 2002; Flores-Lagunes et al., 2010; Flores et al., 2012).

Our empirical results characterize the heterogeneous impact of JC training at different points of the wage distribution. The estimated bounds for a sample that excludes the group of Hispanics suggest positive effects of JC on wages, both at the mean and throughout the wage distribution. For the various non-Hispanic demographic groups analyzed, the statistically significant estimated average effects are bounded between 4.6 and 12 percent, while the statistically significant quantile treatment effects are bounded between 2.7 and 14 percent. Our analysis by race and gender reveals that the positive effects for Blacks appear larger in the lower half of their wage distribution, while for Whites the effects appear larger in the upper half of their wage distribution. Non-Hispanic Females in the lower part of their wage distribution do not show statistically significant positive effects of JC on their wages, while those in the upper part do. Lastly, our set of estimated bounds for Hispanics are wide and include zero.[3]

The paper is organized as follows. Section 2 presents the sample selection problem and the Horowitz and Manski (2000) bounds. Sections 3 and 4 discuss, respectively, bounds on average and quantile treatment effects, as well as the additional assumptions we consider. Section 5 describes the JC program and the NJCS, and Section 6 presents the empirical results from our application. Section 7 concludes.

2 Sample Selection and the Horowitz-Manski Bounds

We describe the sample selection problem in the context of estimating the causal effect of a training program (e.g., JC) on wages, where the problem arises because—even in the presence of random assignment—only the wages of those employed are observed.

[3] Our set of estimated bounds for Hispanics does not employ one of our assumptions (individual-level monotonicity of the treatment on employment) for reasons that will be discussed in subsequent sections.

Formally, consider having access to data on N individuals and define a binary treatment T_i, which indicates whether individual i has participated in the program (T_i=1) or not (T_i=0). We start with the following assumption:

Assumption A. T_i *is randomly assigned.*

To illustrate the sample selection problem, assume for the moment that the individual's wage is a linear function of a constant term, the treatment indicator T_i and a set of pre-treatment characteristics X_{1i},[4]

$$Y_i^* = \beta_0 + T_i\beta_1 + X_{1i}\beta_2 + U_{1i}, \tag{1}$$

where Y_i^* is the latent wage for individual i, which is observed conditional on the self-selection process into employment. This process is also assumed (for the moment) to be linearly related to a constant, the treatment indicator T_i and a set of pre-treatment characteristics X_{2i},

$$S_i^* = \delta_0 + T_i\delta_1 + X_{2i}\delta_2 + U_{2i}. \tag{2}$$

Similarly, S_i^* is a latent variable representing the individual's propensity to be employed. Let S_i denote the observed employment indicator that takes values S_i=1 if individual i is employed and 0 otherwise. Then, $S_i = 1[S_i^* \geq 0]$, where $1[\cdot]$ is an indicator function. Therefore, we observe individual i's wage, Y_i, when i is employed (S_i =1) and it remains latent when unemployed (S_i =0). In this setting, which assumes treatment effects are constant over the population, the parameter of interest is β_1.

Conventionally, point identification of β_1 requires strong assumptions such as joint independence of the errors (U_{1i}, U_{2i}) in the wage and employment equations and the regressors T_i, X_{1i} and X_{2i}, plus bivariate normality of (U_{1i}, U_{2i}) (Heckman, 1979). The bivariate normality assumption about the errors can be relaxed by relying on exclusion restrictions (Heckman, 1990; Heckman and Smith, 1995; Imbens and Angrist, 1994), which require variables that determine employment but do not affect wages, or equivalently, variables in X_{2i} that do not belong in X_{1i}. However, it is well known that finding such variables that go along with economic reasoning in this situation is extremely difficult (Angrist and Krueger, 1999; Angrist and Krueger, 2001). More generally, in many economic applications it is difficult to find valid exclusion restrictions.

[4]Linearity is assumed here to simplify the exposition of the sample selection problem. The non-parametric approach to address sample selection employed in this paper does not impose linearity or functional form assumptions to partially identify the treatment effects of interest.

An alternative approach suggests that the parameters can be bounded without relying on distributional assumptions or on the availability and validity of exclusion restrictions. Horowitz and Manski (2000; HM hereafter) proposed a general framework to construct bounds on treatment effects when data are missing due to a nonrandom process, such as self-selection into non-employment ($S_i^* < 0$), provided that the outcome variable has a bounded support.

To illustrate HM's bounds, let $Y_i(0)$ and $Y_i(1)$ be the potential (counterfactual) wages for unit i under control (T_i=0) and treatment (T_i=1), respectively. The relationship between these potential wages and the observed Y_i is that $Y_i = Y_i(1)T_i + Y_i(0)(1 - T_i)$. Define the average treatment effect (ATE) as:

$$ATE = E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]. \tag{3}$$

Conditional on T_i and the observed employment indicator S_i, the ATE in (3) can be written as:

(4)

$$\begin{aligned} ATE = {} & E[Y_i|T_i = 1, S_i = 1]Pr(S_i = 1|T_i = 1) + E[Y_i(1)|T_i = 1, S_i = 0]Pr(S_i = 0|T_i = 1) \\ & - E[Y_i|T_i = 0, S_i = 1]Pr(S_i = 1|T_i = 0) - E[Y_i(0)|T_i = 0, S_i = 0]Pr(S_i = 0|T_i = 0). \end{aligned}$$

Examination of Equation (4) reveals that, under random assignment, we can identify from the data all the conditional probabilities ($Pr(S_i = s|T_i = t)$, for $(t, s) = (0, 1)$) and also the expectations of the wage when conditioning on S_i=1 ($E[Y_i|T_i = 1, S_i = 1]$ and $E[Y_i|T_i = 0, S_i = 1]$). Sample selection into non-employment makes it impossible to point identify $E[Y_i(1)|T_i = 1, S_i = 0]$ and $E[Y_i(0)|T_i = 0, S_i = 0]$. We can, however, construct HM bounds on these unobserved objects provided that the support of the outcome lies in a bounded interval $[Y^{LB}, Y^{UB}]$, since this implies that the values for these unobserved objects are restricted to such interval. Thus, HM's lower and upper bounds (LB^{HM} and UB^{HM}, respectively) are identified as follows:

$$\begin{aligned} LB^{HM} &= E[Y_i|T_i = 1, S_i = 1]Pr(S_i = 1|T_i = 1) + Y^{LB}Pr(S_i = 0|T_i = 1) \\ &\quad - E[Y_i|T_i = 0, S_i = 1]Pr(S_i = 1|T_i = 0) - Y^{UB}Pr(S_i = 0|T_i = 0) \\ UB^{HM} &= E[Y_i|T_i = 1, S_i = 1]Pr(S_i = 1|T_i = 1) + Y^{UB}Pr(S_i = 0|T_i = 1) \\ &\quad - E[Y_i|T_i = 0, S_i = 1]Pr(S_i = 1|T_i = 0) - Y^{LB}Pr(S_i = 0|T_i = 0) \end{aligned} \tag{5}$$

Note that these bounds do not employ distributional or exclusion restrictions assumptions. They are nonparametric and allow for heterogeneous treatment effects, that is, non-constant effects over the population. On the other hand, a cost of imposing only Assumption A and boundedness of the outcome is that the HM bounds are often wide. Indeed, this is the case in our application, as will be shown below. For this reason, we take this approach as a building block and proceed by imposing more structure through the use of assumptions that are typically weaker than the distributional and exclusion restriction assumptions needed for point identification.

3 Bounds on Average Treatment Effects

Lee (2009) and Zhang et al. (2008) employ monotonicity assumptions that lead to a trimming procedure that tightens the HM bounds. They implicitly or explicitly employ the principal stratification framework of Frangakis and Rubin (2002) to motivate and derive their results. Principal stratification provides a framework for analyzing causal effects when controlling for a post-treatment variable that has been affected by treatment assignment. In the context of analyzing the effect of JC on wages, the affected post-treatment variable is employment. In this framework, individuals are classified into "principal strata" based on the potential values of employment under each treatment arm. Comparisons of outcomes by treatment assignment within strata can be interpreted as causal effects because which strata an individual belongs to is not affected by treatment assignment.

More formally, let the potential values of employment be denoted by $S_i(0)$ and $S_i(1)$ when i is assigned to control and treatment, respectively. We can partition the population into strata based on the values of the vector $\{S_i(0), S_i(1)\}$. Since both S_i and T_i are binary, there are four principal strata defined as $NN : \{S_i(0) = 0, S_i(1) = 0\}$, $EE : \{S_i(0) = 1, S_i(1) = 1\}$, $EN : \{S_i(0) = 1, S_i(1) = 0\}$, and $NE : \{S_i(0) = 0, S_i(1) = 1\}$. In the context of the application to JC, NN is the stratum of those individuals who would not be employed regardless of treatment assignment, while EE is the stratum of those who would be employed regardless of treatment assignment. The stratum EN represents those who would be employed if assigned to control but unemployed if assigned to treatment, and NE is the stratum of those who would be unemployed if assigned to control but employed if assigned to treatment. Given that strata are defined based on the potential values of S_i, the stratum an individual belongs to is unobserved. A mapping of the observed groups based on (T_i, S_i) to the unobserved strata above is depicted in the first two columns of Table 1.

Lee (2009) and Zhang et al. (2008) focus on the average effect of a program on wages for individuals who would be employed regardless of treatment status, i.e., the EE stratum. This stratum is the only one for which wages are observed under both treatment arms, and thus fewer assumptions are required to construct bounds for its effects. The average treatment effect for this stratum is:

$$ATE_{EE} = E[Y_i(1)|EE] - E[Y_i(0)|EE]. \tag{6}$$

3.1 Bounds Adding Individual-Level Monotonicity

To tighten the HM bounds, we employ the following individual-level monotonicity assumption about the relationship between the treatment (JC) and the selection indicator (employment):

Assumption B. *Individual-Level Positive Weak Monotonicity of S in T:* $S_i(1) \geq S_i(0)$ for all i.

This assumption states that treatment assignment weakly affects selection in one direction, effectively ruling out the EN stratum. Both Lee (2009) and Zhang et al. (2008) employed this assumption, and similar assumptions are widely used in the instrumental variable (Imbens and Angrist, 1994) and partial identification literatures (Manski and Pepper, 2000; Bhattacharya et al., 2008; Flores and Flores-Lagunes, 2010). Although Assumption B is directly untestable, Assumptions A and B imply (but are not implied by) $E(S_i|T_i = 1) - E(S_i|T_i = 0) \geq 0$, which provides a testable implication for Assumption B (Imai, 2008) in settings where Assumption A holds by design, as in our application. In other words, if the testable implication is not satisfied, then Assumption B is not consistent with the data. If it is satisfied, it only means that the data is consistent with this particular testable implication, however, it does not imply that Assumption B is valid. Thus, this testable implication is not to be interpreted as a statistical test of Assumption B.

In the context of JC, Assumption B seems plausible since one of the program's stated goals is to increase the employability of participants. It does so by providing academic, vocational and social skills training to participants, as well as job search assistance. Indeed, the NJCS reported a positive and highly statistically significant average effect of JC on employment of four percentage points (Schochet et al., 2001). Nevertheless, since this assumption is imposed at the individual level, it may be hard to satisfy as it requires that no individual has a negative effect of the program on employment (or, more generally, on selection).

Two factors that may cast doubt on this assumption in our application are that individuals are "locked-in" away from employment while undergoing training (van Ours, 2004), and the possibility that trained individuals may have a higher reservation wage after training and thus may choose to remain unemployed (e.g., Blundell et al., 2007). Note, however, that these two factors become less relevant the longer the time horizon after randomization at which the outcome is measured. For this reason, in Section 6 we focus on wages at the 208^{th} week after random assignment, which is the latest wage measure available in the NJCS.[5] In addition, there is one demographic group in our sample for which Assumption B is more likely to be violated. Hispanics in the NJCS were the only group found to have negative but statistically insignificant mean effects of JC on earnings and employment of -\$15.1 and -3.1, respectively (Schochet et al., 2001; Flores-Lagunes et al., 2010). Although this does not show that the testable implication of Assumption B is statistically rejected for Hispanics, it casts doubt on the validity of this assumption for this group. Thus, we conduct a separate analysis for Hispanics that does not employ Assumption B in Section 6.

Assumption B, by virtue of eliminating the EN stratum, allows the identification of some individuals in the EE and NN strata, as can be seen after deleting the EN stratum in the last column of Table 1. Furthermore, the combination of Assumptions A and B point identifies the proportions of each principal stratum in the population. Let π_k be the population proportions of each principal stratum, $k = NN, EE, EN, NE$, and let $p_{s|t} \equiv Pr(S_i = s|T_i = t)$ for $t, s = 0, 1$. Then, $\pi_{EE} = p_{1|0}, \pi_{NN} = p_{0|1}, \pi_{NE} = p_{1|1} - p_{1|0} = p_{0|0} - p_{0|1}$ and $\pi_{EN} = 0$. Looking at the last column of Table 1, we know that individuals in the observed group with $(T_i, S_i) = (0, 1)$ belong to the stratum of interest EE. Therefore, we can point identify $E[Y_i(0)|EE]$ in (6) as $E[Y_i|T_i = 0, S_i = 1]$. However, it is not possible to point identify $E[Y_i(1)|EE]$, since the observed group with

[5] Zhang et al. (2009) provide some evidence that the estimated proportion of individuals who do not satisfy the individual-level assumption (the EN stratum) falls with the time horizon at which the outcome is measured after randomization.

$(T_i, S_i) = (1,1)$ is a mixture of individuals from two strata, EE and NE. Nevertheless, it can be bounded. Write $E[Y_i|T_i = 1, S_i = 1]$ as a weighted average of individuals belonging to the EE and NE strata:

$$E[Y_i|T_i = 1, S_i = 1] = \frac{\pi_{EE}}{(\pi_{EE} + \pi_{NE})} E[Y_i(1)|EE] + \frac{\pi_{NE}}{(\pi_{EE} + \pi_{NE})} E[Y_i(1)|NE] \tag{7}$$

Since the proportion of EE individuals in the group $(T_i, S_i) = (1,1)$ can be point identified as $\pi_{EE}/(\pi_{EE}+\pi_{NE})$=$p_{1|0}/p_{1|1}$, $E[Y_i(1)|EE]$ can be bounded from above by the expected value of Y_i for the $(p_{1|0}/p_{1|1})$ fraction of the largest values of Y_i in the observed group (T_i, S_i)=(1, 1). In other words, the upper bound is obtained under the scenario that the largest $(p_{1|0}/p_{1|1})$ values of Y_i belong to the EE individuals. Thus, computing the expected value of Y_i after trimming the lower tail of the distribution of Y_i in (T_i, S_i)=(1, 1) by $1 - (p_{1|0}/p_{1|1})$ yields an upper bound for the EE group. Similarly, $E[Y_i(1)|EE]$ can be bounded from below by the expected value of Y_i for the $(p_{1|0}/p_{1|1})$ fraction of the smallest values of Y_i for those in the same observed group. The resulting upper (UB_{EE}) and lower (LB_{EE}) bounds for ATE_{EE} are (Lee, 2009; Zhang et al., 2008):

$$\begin{aligned} UB_{EE} &= E[Y_i|T_i = 1, S_i = 1, Y_i \geq y^{11}_{1-(p_{1|0}/p_{1|1})}] - E[Y_i|T_i = 0, S_i = 1] \\ LB_{EE} &= E[Y_i|T_i = 1, S_i = 1, Y_i \leq y^{11}_{(p_{1|0}/p_{1|1})}] - E[Y_i|T_i = 0, S_i = 1], \end{aligned} \tag{8}$$

where $y^{11}_{1-(p_{1|0}/p_{1|1})}$ and $y^{11}_{(p_{1|0}/p_{1|1})}$ denote the $1 - (p_{1|0}/p_{1|1})$ and the $(p_{1|0}/p_{1|1})$ quantiles of Y_i conditional on $T_i = 1$ and $S_i = 1$, respectively. Lee (2009) shows that these bounds are sharp (i.e., there are no shorter bounds possible under the current assumptions).

The bounds in (8) can be estimated with sample analogs:

$$\begin{aligned} \widehat{UB_{EE}} &= \frac{\Sigma^n_{i=1} Y_i \cdot T_i \cdot S_i \cdot 1[Y_i \geq \widehat{y_{1-\hat{p}}}]}{\Sigma^n_{i=1} T_i \cdot S_i \cdot 1[Y_i \geq \widehat{y_{1-\hat{p}}}]} - \frac{\Sigma^n_{i=1} Y_i \cdot (1 - T_i) \cdot S_i}{\Sigma^n_{i=1} (1 - T_i) \cdot S_i} \\ \widehat{LB_{EE}} &= \frac{\Sigma^n_{i=1} Y_i \cdot T_i \cdot S_i \cdot 1[Y_i \leq \widehat{y_{\hat{p}}}]}{\Sigma^n_{i=1} T_i \cdot S_i \cdot 1[Y_i \leq \widehat{y_{\hat{p}}}]} - \frac{\Sigma^n_{i=1} Y_i \cdot (1 - T_i) \cdot S_i}{\Sigma^n_{i=1} (1 - T_i) \cdot S_i}, \end{aligned} \tag{9}$$

where $\widehat{y_{1-\hat{p}}}$ and $\widehat{y_{\hat{p}}}$ are the sample analogs of the quantities $y^{11}_{1-(p_{1|0}/p_{1|1})}$ and $y^{11}_{(p_{1|0}/p_{1|1})}$ in (8), respectively, and $\hat{p}$, the sample analog of $(p_{1|0}/p_{1|1})$, is calculated as follows:

$$\hat{p} = \frac{\Sigma^n_{i=1} (1 - T_i) \cdot S_i}{\Sigma^n_{i=1} (1 - T_i)} / \frac{\Sigma^n_{i=1} T_i \cdot S_i}{\Sigma^n_{i=1} T_i}. \tag{10}$$

Lee (2009) shows that these estimators are asymptotically normal.

3.2 Tightening the Bounds by Adding Weak Monotonicity of Mean Potential Outcomes Across Strata

We present a weak monotonicity assumption of mean potential outcomes across the EE and NE strata that tightens the bounds in (8). This assumption was originally proposed by Zhang and Rubin (2003) and employed in Zhang et al. (2008):

Assumption C. *Weak Monotonicity of Mean Potential Outcomes Across the EE and NE Strata:* $E[Y(1)|EE] \geq E[Y(1)|NE]$.

Intuitively, in the context of JC, this assumption formalizes the notion that the EE stratum is likely to be comprised of more "able" individuals than those belonging to the NE stratum. Since "ability" is positively correlated with labor market outcomes, one would expect wages for the individuals who are employed regardless of treatment status (the EE stratum) to weakly dominate on average the wages of those individuals who are employed only if they receive training (the NE stratum). Hence, Assumption C requires a positive correlation between employment and wages. While Assumption C is not directly testable, one can indirectly gauge its plausibility by comparing the average of pre-treatment covariates that are highly correlated with wages between the EE and NE strata, as we illustrate in Section 6.2 below.

Employing Assumptions A, B, and C results in tighter bounds. To see this, recall that the average outcome in the observed group with $(T_i, S_i) = (1, 1)$ contains units from two strata, EE and NE, and can be written as the weighted average shown in (7). By replacing $E[Y_i(1)|NE]$ with $E[Y_i(1)|EE]$ in (7) and using the inequality in Assumption C, we have that $E[Y_i|T_i = 1, S_i = 1] \leq E[Y_i(1)|EE]$, and thus that $E[Y_i(1)|EE]$ is bounded from below by $E[Y_i|T_i = 1, S_i = 1]$. Therefore, the lower bound for ATE_{EE} becomes: $E[Y_i|T_i = 1, S_i = 1] - E[Y_i|T_i = 0, S_i = 1]$. Imai (2008) shows that these bounds are sharp.

To estimate the bounds under Assumptions A, B, and C, note that the upper bound estimator of (8) remains $\widehat{UB_{EE}}$ from (9), while the estimator of the lower bound is the corresponding sample analog of $E[Y_i|T_i = 1, S_i = 1] - E[Y_i|T_i = 0, S_i = 1]$:

$$\widehat{LB^c_{EE}} = \frac{\Sigma_{i=1}^n Y_i \cdot T_i \cdot S_i}{\Sigma_{i=1}^n T_i \cdot S_i} - \frac{\Sigma_{i=1}^n Y_i \cdot (1 - T_i) \cdot S_i}{\Sigma_{i=1}^n (1 - T_i) \cdot S_i}. \tag{11}$$

3.3 Narrowing Bounds on ATE_{EE} using a Covariate

Under Assumptions A and B, Lee (2009) shows that (i) grouping the sample based on the values of a pre-treatment covariate X, (ii) applying the trimming procedure to construct bounds for each group, and (iii) averaging the bounds across these groups, results in narrower bounds for ATE_{EE} as compared to those in (8). This result follows from the properties of trimmed means, and thus it is applicable only to bounds that involve trimming.[6]

Let X take values on $\{x_1, \ldots, x_J\}$. By the law of iterated expectations, we can write the non-point identified term in (6) as:

$$E[Y_i(1)|EE] = E_X\{E[Y_i(1)|EE, X_i = x_j]|EE\}. \tag{12}$$

Recall from (8) that the bounds on $E[Y_i(1)|EE]$ without employing X are given by $E[Y_i|T_i = 1, S_i = 1, Y_i \geq y^{11}_{1-(p_{1|0}/p_{1|1})}] \geq E[Y_i(1)|EE] \geq E[Y_i|T_i = 1, S_i = 1, Y_i \leq y^{11}_{(p_{1|0}/p_{1|1})}]$. Thus, it is straightforward to construct bounds on the terms $E[Y_i(1)|EE, X_i = x_j]$ for the different values of X by implementing the trimming bounds on $E[Y_i(1)|EE]$ discussed in Sections 3.1 *within* cells with $X_i = x_j$. Let these bounds be denoted by $LB^{Y(1)}_{EE}(x_j)$ and $UB^{Y(1)}_{EE}(x_j)$, so that $UB^{Y(1)}_{EE}(x_j) \geq E[Y_i(1)|EE, X_i = x_j] \geq LB^{Y(1)}_{EE}(x_j)$. It is important to note that the trimming proportions will differ across groups with different values of X, as the conditional probabilities $p_{s|t}$ are now computed within cells with $X_i = x_j$. After substituting the trimming bounds on $E[Y_i(1)|EE, X_i = x_j]$ into equation (12) we obtain the bounds on ATE_{EE}, which are given by

$$\begin{aligned} UB^*_{EE} &= E_X\{UB^{Y(1)}_{EE}(x_j)|EE\} - E[Y_i|T_i = 0, S_i = 1] \\ LB^*_{EE} &= E_X\{LB^{Y(1)}_{EE}(x_j)|EE\} - E[Y_i|T_i = 0, S_i = 1]. \end{aligned} \tag{13}$$

Lee (2009) shows that, under Assumptions A and B, these bounds are sharp and that, as compared to those in (8), $UB^*_{EE} \leq UB_{EE}$ and $LB^*_{EE} \geq LB_{EE}$.

An important step in the computation of the bounds in (13) is the estimation of the term $\Pr(X = x_j|EE)$ used in computing the outer expectation in the first term. By Bayes' rule, we can write $\Pr(X = x_j|EE) = \pi_{EE}(x_j) \cdot \Pr(X = x_j) / \left[\sum_{j=1}^{J} \pi_{EE}(x_j) \cdot \Pr(X = x_j)\right]$,

[6] The key property is that the mean of a lower (upper) tail truncated distribution is greater (less) than or equal to the average of the means of lower (upper) tail truncated distributions conditional on $X = x$, given that the proportion of the overall population that is eventually trimmed is the same.

where $\pi_{EE}(x_j) = \Pr(EE|X = x_j)$ is the EE stratum proportion in the cell $X = x_j$. Thus, the sample analog estimators of the bounds in (13) are:

$$\begin{aligned}
\widehat{UB}^{*}_{EE} &= \sum_{j=1}^{J} \widehat{UB}^{Y(1)}_{EE}(x_j)\widehat{\Pr}(X = x_j|EE) - \frac{\sum_{i=1}^{n} Y_i \cdot (1 - T_i) \cdot S_i}{\sum_{i=1}^{n}(1 - T_i) \cdot S_i} \\
\widehat{LB}^{*}_{EE} &= \sum_{j=1}^{J} \widehat{LB}^{Y(1)}_{EE}(x_j)\widehat{\Pr}(X = x_j|EE) - \frac{\sum_{i=1}^{n} Y_i \cdot (1 - T_i) \cdot S_i}{\sum_{i=1}^{n}(1 - T_i) \cdot S_i},
\end{aligned}$$

where $\widehat{UB}^{Y(1)}_{EE}(x_j)$ and $\widehat{LB}^{Y(1)}_{EE}(x_j)$ are the estimators of $UB^{Y(1)}_{EE}(x_j)$ and $LB^{Y(1)}_{EE}(x_j)$, respectively, which are computed using the estimators of the bounds on $E[Y_i(1)|EE]$ in the first term of (9) for individuals with $X_i = x_j$, and

$$\widehat{\Pr}(X = x_j|EE) = \frac{\left[\sum_{i=1}^{n}(1 - T_i)S_i 1\left[X_i = x_j\right] / \sum_{i=1}^{n}(1 - T_i)1\left[X_i = x_j\right]\right]\left[\sum_{i=1}^{n} 1\left[X_i = x_j\right]\right]}{\sum_{j=1}^{J}\left\{\left[\sum_{i=1}^{n}(1 - T_i)S_i 1\left[X_i = x_j\right] / \sum_{i=1}^{n}(1 - T_i)1\left[X_i = x_j\right]\right]\left[\sum_{i=1}^{n} 1\left[X_i = x_j\right]\right]\right\}}.$$

Finally, under Assumptions A, B and C the procedure above is only applied to the upper bound on ATE_{EE}, as the lower bound does not involve trimming.

4 Bounds on Quantile Treatment Effects

We now extend the results presented in the previous section to construct bounds on quantile treatment effects (QTE) based on results by Imai (2008). The parameters of interest are differences in the quantiles of the marginal distributions of the potential outcomes $Y(1)$ and $Y(0)$; more specifically, we define the α-quantile effect for the EE stratum:

$$QTE^{\alpha}_{EE} = F^{-1}_{Y_i(1)|EE}(\alpha) - F^{-1}_{Y_i(0)|EE}(\alpha), \tag{14}$$

where $F^{-1}_{Y_i(t)|EE}(\alpha)$ denotes the α-quantile of the distribution of $Y_i(t)$ for the EE stratum.

Two recent papers have focused on partial identification of QTE. Blundell et. al., (2007) derived sharp bounds on the distribution of wages and the interquantile range to study income inequality in the U.K. Their work builds on the bounds on the conditional quantiles in Manski (1994), which are tightened by imposing stochastic dominance assumptions. Their stochastic dominance assumption is applied to the distribution of wages of individuals observed employed and unemployed, whereby the wages of employed individuals are assumed to weakly dominate those of unemployed individuals (i.e., positive selection into employment). In addition, they explore the use of exclusion restrictions to further tighten their bounds. Lechner and Melly (2010) analyze QTE of a German training program on wages. They impose an individual-level monotonicity assumption similar to our Assumption B, and employ the stochastic dominance assumption of Blundell et al. (2007) to tighten their bounds. In contrast to those papers, we take advantage of the randomization in the NJCS to estimate QTE by employing individual-level monotonicity (Assumption B) and by strengthening Assumption C to stochastic dominance applied to the EE and NE strata. Another difference is the parameters of interest: Blundell et al. (2007) focus on the population QTE, Lechner and Melly (2010) focus on the QTE for those individuals who are employed under treatment, and our focus is on the QTE for individuals who are employed regardless of treatment assignment.[7]

Let $F_{Y_i|T_i=t,S_i=s}(\cdot)$ be the cumulative distribution of individuals' wages conditional on $T_i = t$ and $S_i = s$, and let y_α^{ts} denote its corresponding α-quantile, for $\alpha \in (0,1)$, or $y_\alpha^{ts} = F^{-1}_{Y_i|T_i=t,S_i=s}(\alpha)$. Under Assumptions A and B, we can partially identify QTE^α_{EE} as $LB^\alpha_{EE} \leq QTE^\alpha_{EE} \leq UB^\alpha_{EE}$, where (Imai, 2008):

$$
\begin{aligned}
UB^\alpha_{EE} &= F^{-1}_{Y_i|T_i=1,S_i=1,Y_i \geq y^{11}_{1-(p_{1|0}/p_{1|1})}}(\alpha) - F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha) \\
LB^\alpha_{EE} &= F^{-1}_{Y_i|T_i=1,S_i=1,Y_i \leq y^{11}_{(p_{1|0}/p_{1|1})}}(\alpha) - F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha).
\end{aligned} \tag{15}
$$

The intuition behind this result is the same as that for the bounds on ATE_{EE} in (8). $F^{-1}_{Y_i|T_i=1,S_i=1,Y_i \geq y^{11}_{1-(p_{1|0}/p_{1|1})}}(\alpha)$ and $F^{-1}_{Y_i|T_i=1,S_i=1,Y_i \leq y^{11}_{(p_{1|0}/p_{1|1})}}(\alpha)$ correspond to the α-quantile of Y_i after trimming, respectively, the lower and upper tail of the distribution of Y_i in $(T_i, S_i) = (1,1)$ by $1 - (p_{1|0}/p_{1|1})$, and thus they provide an upper and lower bound for $F^{-1}_{Y_{i(1)}|EE}(\alpha)$ in (14). Similar to (8), the quantile $F^{-1}_{Y_{i(0)}|EE}(\alpha)$ is point identified from the group with $(T_i, S_i) = (1,0)$. Imai (2008) shows that the bounds in (15) are sharp.

Using the same notation as in (9), we estimate the bounds in (15) using their sample

[7]The treated-and-employed subpopulation of Lechner and Melly (2010) is a mixture of two strata: EE and NE. In our application, the EE stratum and the treated-and-employed subpopulation account for about the same proportion of the population (57 and 61 percent, respectively).

analogs:

$$\widehat{UB^{\alpha}_{EE}} = \min\left\{y : \frac{\sum_{i=1}^{n} T_i \cdot S_i \cdot 1\left[Y_i \geq \widehat{y_{1-\widehat{p}}}\right] \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n} T_i \cdot S_i \cdot 1\left[Y_i \geq \widehat{y_{1-\widehat{p}}}\right]} \geq \alpha\right\} - \min\left\{y : \frac{\sum_{i=1}^{n}(1-T_i) \cdot S_i \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n}(1-T_i) \cdot S_i} \geq \alpha\right\}$$

(16)

$$\widehat{LB^{\alpha}_{EE}} = \min\left\{y : \frac{\sum_{i=1}^{n} T_i \cdot S_i \cdot 1\left[Y_i \leq \widehat{y_{\widehat{p}}}\right] \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n} T_i \cdot S_i \cdot 1\left[Y_i \leq \widehat{y_{\widehat{p}}}\right]} \geq \alpha\right\} - \min\left\{y : \frac{\sum_{i=1}^{n}(1-T_i) \cdot S_i \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n}(1-T_i) \cdot S_i} \geq \alpha\right\}.$$

4.1 Tightening Bounds on QTE^{α}_{EE} using Stochastic Dominance

We tighten the bounds in (15) by strengthening Assumption C to stochastic dominance. Let $F_{Y_i(1)|EE}(\cdot)$ and $F_{Y_i(1)|NE}(\cdot)$ denote the cumulative distributions of $Y_i(1)$ for individuals who belong to the EE and NE strata, respectively.:

Assumption D. *Stochastic Dominance Between the EE and NE Strata*: $F_{Y_i(1)|EE}(y) \leq F_{Y_i(1)|NE}(y)$, for all y.

This assumption directly imposes restrictions on the distribution of potential outcomes under treatment for individuals in the EE stratum, which results in a tighter lower bound relative to that in (15). Under Assumptions A, B, and D, the resulting sharp bounds are (Imai, 2008): $LB^{d\alpha}_{EE} \leq QTE^{\alpha}_{EE} \leq UB^{\alpha}_{EE}$, where UB^{α}_{EE} is as in (15) and

$$LB^{d\alpha}_{EE} = F^{-1}_{Y_i|T_i=1,S_i=1}(\alpha) - F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha). \tag{17}$$

The estimator of the upper bound is still given by $\widehat{UB^{\alpha}_{EE}}$ in (16), while the estimator for $LB^{d\alpha}_{EE}$ is now given by:

(18)

$$\widehat{LB^{d\alpha}_{EE}} = \min\left\{y : \frac{\sum_{i=1}^{n} T_i \cdot S_i \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n} T_i \cdot S_i} \geq \alpha\right\} - \min\left\{y : \frac{\sum_{i=1}^{n}(1-T_i) \cdot S_i \cdot 1\left[Y_i \leq y\right]}{\sum_{i=1}^{n}(1-T_i) \cdot S_i} \geq \alpha\right\}$$

4.2 Narrowing Bounds on QTE^{α}_{EE} using a Covariate

In this section we propose a way to use a pre-treatment covariate X taking values on $\{x_1, \ldots, x_J\}$ to narrow the trimming bounds on $F^{-1}_{Y_i(1)|EE}(\alpha)$ and, thus, the bounds on QTE^{α}_{EE} in (15). The idea is similar to that in Lee (2009) described in Section 3.3, however, the non-linear form of the quantile function $F^{-1}_{Y_i(1)|EE}(\alpha)$ prevents us from directly using the law of iterated expectations as in (12). To circumvent this difficulty, we first focus on the cumulative distribution function (CDF) of $Y_i(1)$ for the stratum EE at a given point

$\widetilde{y}$, $F_{Y_i(1)|EE}(\widetilde{y})$, and write it as the mean of an indicator function, which allows us to use iterated expectations. A similar approach was also used in Lechner and Melly (2010) to control for selection into treatment based on covariates. Using this insight we can write:

$$F_{Y_i(1)|EE}(\widetilde{y}) = E[1[Y_i(1) \leq \widetilde{y}]|EE] = E_X\{E[1[Y_i(1) \leq \widetilde{y}]|EE, X_i = x_j]|EE\}. \tag{19}$$

Note that (19) is similar to (12) except that we now employ $1[Y_i(1) \leq \widetilde{y}]$ as the outcome instead of $Y_i(1)$. Thus, the methods discussed in Section 3.3 (and more generally, the trimming bounds in Section 3.1) can be used to bound $F_{Y_i(1)|EE}(\widetilde{y})$. As in Section 3.3, let $UB^{\widetilde{y}}_{EE}(x_j)$ and $LB^{\widetilde{y}}_{EE}(x_j)$ denote the upper and lower bounds on $E[1[Y_i(1) \leq \widetilde{y}]|EE, X_i = x_j]$ under Assumptions A and B, which are just the trimming bounds on $E[Y_i(1)|EE]$ in the first part of (8) within cells with $X_i = x_j$ and employing as outcome the indicator $1(Y_i \leq \widetilde{y})$ instead of Y_i. After substituting $UB^{\widetilde{y}}_{EE}(x_j)$ and $LB^{\widetilde{y}}_{EE}(x_j)$ into (19) we obtain the following upper and lower bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ under Assumptions A and B:

$$\begin{aligned} F_{UB}(\widetilde{y}) &= E_X\{UB^{\widetilde{y}}_{EE}(x_j)|EE\} \\ F_{LB}(\widetilde{y}) &= E_X\{LB^{\widetilde{y}}_{EE}(x_j)|EE\}. \end{aligned} \tag{20}$$

Importantly, by the results in Lee (2009) discussed in Section 3.3, these trimming bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ are sharp and tighter than those not employing the covariate X.

Given bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ for all $\widetilde{y} \in \Re$, the lower (upper) bound on the α-quantile of $Y_i(1)$ for the EE stratum, $F^{-1}_{Y_i(1)|EE}(\alpha)$, is obtained by inverting the upper (lower) bound on $F_{Y_i(1)|EE}(\widetilde{y})$. Using the bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ in (20), the lower and upper bounds on $F^{-1}_{Y_i(1)|EE}(\alpha)$ are obtained by finding the value y_α such that $F_{UB}(y_\alpha) = \alpha$ and $F_{LB}(y_\alpha) = \alpha$, respectively.[8] Therefore, the bounds on QTE^{α}_{EE} under Assumptions A and B are given by:

$$\begin{aligned} UB^{*\alpha}_{EE} &= F^{-1}_{LB}(\alpha) - F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha) \\ LB^{*\alpha}_{EE} &= F^{-1}_{UB}(\alpha) - F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha). \end{aligned} \tag{21}$$

We implement this procedure by estimating the bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ in (20) at M different values of $\widetilde{y}$ spanning the support of the outcome, and then inverting the resulting estimated bounds to obtain the estimate of the bounds on the α-quantile $F^{-1}_{Y_i(1)|EE}(\alpha)$. This last set of estimated bounds are then combined with the estimate of $F^{-1}_{Y_i|T_i=0,S_i=1}(\alpha)$

[8]Note that, since we are inverting the CDF, the lower (upper) bound on the quantile is computed employing the upper (lower) bound of the CDF.

to compute estimates of the bounds on QTE^{α}_{EE} in (21). The bounds on $F_{Y_i(1)|EE}(\widetilde{y}_m)$ at each point $\widetilde{y}_m$ $(m = 1, \ldots, M)$ in (20) are estimated employing the estimators of the bounds on $E[Y_i(1)|EE]$ in the first term of (9) for individuals with $X_i = x_j$ and using as outcome the indicator function $1(Y_i \leq \widetilde{y}_m)$ instead of Y_i. Finally, just as in the case of the ATE_{EE} in Section 3.3, under Assumptions A, B and D, the procedure above is only applied to the upper bound because the lower bound does not involve trimming.

5 Job Corps and the National Job Corps Study

We employ the methods described in the previous sections to assess the effect of Job Corps (JC) on the wages of participants. JC is America's largest and most comprehensive education and job training program. It is federally funded and currently administered by the US Department of Labor. With a yearly cost of about \$1.5 billion, JC annual enrollment ascends to 100,000 students (US Department of Labor, 2010). The program's goal is to help disadvantaged young people, ages 16 to 24, improve the quality of their lives by enhancing their labor market opportunities and educational skills set. Eligible participants receive academic, vocational, and social skills training at over 123 centers nationwide (US Department of Labor, 2010), where they typically reside. Participants are selected based on several criteria, including age, legal US residency, economically disadvantage status, living in a disruptive environment, in need of additional education or training, and be judged to have the capability and aspirations to participate in JC (Schochet et al., 2001).

Being the nation's largest job training program, the effectiveness of JC has been debated at times. During the mid nineties, the US Department of Labor funded the National Job Corps Study (NJCS) to determine the program's effectiveness. The main feature of the study was its random assignment: individuals were taken from nearly all JC's outreach and admissions agencies located in the 48 continuous states and the District of Columbia and randomly assigned to treatment and control groups. From a randomly selected research sample of 15,386 first time eligible applicants, approximately 61 percent were assigned to the treatment group (9,409) and 39 percent to the control group (5,977), during the sample intake period from November 1994 to February 1996. After recording their data through a baseline interview for both treatment and control experimental groups, a series of follow up interviews were conducted at weeks 52, 130, and 208 after randomization (Schochet et al., 2001).

Randomization took place before participants' assignment to a JC center. As a result, only 73 percent of the individuals randomly assigned to the treatment group actually enrolled in JC. Also, about 1.4 percent of the individuals assigned to the control group enrolled in the program despite the three-year embargo imposed on them (Schochet et al., 2001). Therefore, in the presence of this non-compliance, the comparison of outcomes by random assignment to the treatment has the interpretation of the "intention-to-treat" (ITT) effect, that is, the causal effect of being offered participation in JC. Focusing on this parameter in the presence of non-compliance is common practice in the literature (e.g., Lee, 2009; Flores-Lagunes et al., 2010; Zhang et al., 2009). Correspondingly, our empirical analysis estimates informative non-parametric bounds for ITT effects, although for simplicity we describe our results in the context of treatment effects.

Our sample is restricted to individuals who have non-missing values for weekly earnings and weekly hours worked for every week after random assignment, resulting in a sample size of 9,145.[9] This is the same sample employed by Lee (2009), which facilitates comparing the informational content of our additional assumption (Assumption C) to tighten the estimated bounds. We also analyze the wage effects of JC for the following demographic groups: Non-Hispanics, Blacks, Whites, Non-Hispanic Males, Non-Hispanic Females, and Hispanics. As we further discuss in Section 6.1, we separate Hispanics in order to increase the likelihood that Assumption B holds. Finally, we employ the NJCS design weights (Schochet, 2001) throughout the analysis, since different subgroups in the population had different probabilities of being included in the research sample.

A potential concern with the NJCS data is measurement error (ME) in the variables of interest (wages, employment, and random assignment) and the extent to which it may affect our estimated bounds. While random assignment (T) is likely to be measured without error, both employment (S) and wages (Y) are self-reported and thus more likely to suffer from this problem. In principle, it is hard to know a priori the effect of ME on the estimated bounds, although accounting for plausible forms of ME will likely lead to wider bounds.[10] Note that ME in S may lead to misclassification of individuals into

[9] As a consequence, we implicitly assume—as do the studies cited in the previous paragraph—that the missing values are "missing completely at random".

[10] We are unaware of work assessing the effect of ME on estimated bounds that do not account for this

strata, affecting the trimming proportions of the bounds involving trimming; while ME in Y will likely affect the trimmed distributions and their moments.

Summary statistics for the sample of 9,145 individuals, which essentially replicate those of Lee (2009, p. 1075), are presented in the Internet Appendix. Pre-treatment variables in the data include: demographic variables, education and background variables, income variables, and employment information. As expected, given the randomization, the distribution of these pre-treatment characteristics is similar between treatment and control groups, with the difference in the means of both groups being not statistically significant at a 5 percent level. The corresponding differences in labor market outcomes at week 208 after randomization for this sample is quantitatively equivalent and consistent with the previously found 12 percent positive effect of JC on participants' weekly earnings and the positive effect on employment of 4 percentage points (Schochet et al., 2001). The effect of JC on participants' weekly hours worked in our sample of about two hours a week is also consistent with that obtained in Schochet et al. (2001). Similar summary statistics for the demographic groups to be analyzed are also relegated to the Internet Appendix.

6 Bounds on the Effect of Job Corps on Wages

We start by presenting the HM bounds, which are the basis for the other bounds discussed in Sections 3 and 4. To construct bounds on the average treatment effect of JC on wages, the HM bounds combine the random assignment in the NJCS (Assumption A) with the empirical bounds of the outcome (log wages at week 208 after randomization). The empirical upper bound on the support of log wages at week 208, denoted by Y^{UB} in (5), is 5.99; while the lower bound, Y^{LB}, is -1.55. Using the expressions in (5), the HM bounds are $UB^{HM} = 3.135$ and $LB^{HM} = -3.109$, with a width of 6.244. Clearly, these bounds are wide and largely uninformative. In what follows, we add assumptions to tighten them.

---

feature of the data. A growing literature that employs bounding techniques to *deal with* ME includes Horowitz and Manski (1995), Bollinger (1996), Molinari (2008) and Gundersen, Kreider, and Pepper (2012). Extending the bounds in this paper to account for ME is beyond its scope.

6.1 Bounds on ATE_{EE} Adding Individual-Level Monotonicity

Under individual-level monotonicity of JC on employment (Assumption B) we partially identify the average effect of JC on wages for those individuals who are employed regardless of treatment assignment (the EE stratum). Therefore, it is of interest to estimate the size of that stratum relative to the full population, which can be done under Assumptions A and B. Table 2 reports the estimated strata proportions for the Full Sample and for the demographic groups we consider. The EE stratum accounts for close to 57 percent of the population, making it the largest stratum. The second largest stratum is the "never employed" or NN, accounting for 39 percent of the population. Lastly, the NE stratum accounts for 4 percent (the stratum EN is ruled out by Assumption B). The relative magnitudes of the strata largely hold for all demographic groups (except NE for Hispanics). Interestingly, Whites have the highest proportion of EE individuals at 66 percent, while Blacks have the lowest at 51 percent.

The second column in Panel A of Table 3 reports estimated bounds for ATE_{EE} for the Full Sample using (9) under Assumptions A and B.[11] Relative to the HM bounds, these bounds are much tighter: their width goes from 6.244 in the HM bounds to 0.121. Unlike the HM bounds, the present bounding procedure does not depend on the empirical support of the outcome. However, the bounds still include zero, as does the Imbens and Manski (2004; IM hereafter) confidence intervals reported in the last row of the panel. These confidence intervals include the true parameter of interest with a 95 percent probability. Thus, while Assumption B greatly tightens the HM bounds, it is not enough to rule out zero or a small negative effect of JC on log wages at week 208.

As discussed in Section 3.1, the untestable individual-level weak monotonicity assumption of the effect of JC on employment may be dubious in certain circumstances. In the context of JC, it has been documented that Hispanics in the NJCS exhibited negative (albeit not statistically significant) average effects of JC on both their employment and weekly earnings, while for the other groups these effects were positive and highly statistically significant (Schochet et al., 2001; Flores-Lagunes et al., 2010). Although this evidence does not show that the testable implication of Assumption B discussed in Section 3.1 is statistically rejected for Hispanics, it casts doubt on the validity of As-

[11] Note that the results for the Full Sample reported in Panel A of Table 3 do not replicate those in Lee (2009) since he employs a transformation of the reported wages (see footnote 13 in that paper).

sumption B for this group. Therefore, we consider a sample that excludes Hispanics, as well as other Non-Hispanic demographic groups (Whites, Blacks, Non-Hispanic Males, and Non-Hispanic Females). We defer the analysis of Hispanics that does not employ Assumption B to Section 6.7, where we also discuss other features of this group in the NJCS.

The remaining columns in Panel A of Table 3 present estimated bounds under Assumptions A and B for various demographic groups, along with their width and 95 percent IM confidence intervals. The third column presents the corresponding estimated bounds for the Non-Hispanics sample. The upper bound for this group is larger than the one for the Full Sample, while the lower bound is less negative, which is consistent with the discussion above regarding Hispanics. The IM confidence intervals are wider for the Non-Hispanics sample relative to the Full Sample, but they are more concentrated on the positive side of the real line. For the other groups (Whites, Blacks, and Non-Hispanic Females and Males), none of the estimated bounds exclude zero, although Whites and Non-Hispanic Males have a lower bound almost right at zero. In general, the IM confidence intervals for the last four demographic groups are wider than those of the Full Sample and Non-Hispanics groups, which is a consequence of their smaller sample sizes.

We now check the testable implication of Assumption B mentioned in Section 3.1: $E(S_i|T_i = 1) - E(S_i|T_i = 0) \geq 0$. The left-hand-side of this expression is the proportion of individuals in the NE stratum (π_{NE}), which is reported in Table 2 for all groups. From that table, it can be seen that for all Non-Hispanic groups the estimated NE stratum proportions are between 0.04 and 0.06, and they are statistically significant at a 1 percent level (not shown in the table). For Hispanics, however, the corresponding proportion is a statistically insignificant 0.002. Thus, while the testable implication of Assumption B is strongly satisfied for all Non-Hispanic groups, the data does not provide evidence in favor of it for Hispanics, making Assumption B dubious for this demographic group.

We close this section by noting, as does Lee (2009), that small and negative estimated lower bounds on the effect of JC on wages under the current assumptions can be interpreted as pointing toward positive effects. The reason is that the lower bound is obtained by placing individuals in the EE stratum at the bottom of the outcome distribution of the observed group with $(T_i, S_i) = (1, 1)$. While this mathematically identifies a valid lower bound, it implies a perfect negative correlation between employment and wages that is implausible from the standpoint of standard models of labor supply, in which individuals with higher predicted wages are more likely to be employed. Indeed, one interpretation that can be given to Assumption C (employed in the next section) is that of formalizing this theoretical notion to tighten the lower bound.

6.2 Bounds on ATE_{EE} Adding Weak Monotonicity of Mean Potential Outcomes Across Strata

Panel B of Table 3 presents the estimated bounds adding Assumption C for the Full Sample (second column) and the Non-Hispanic demographic groups. This assumption has considerable identifying power as it results in tighter bounds for the ATE_{EE} compared to the previously estimated bounds, with the width being cut in about half for the Full Sample. Importantly, employing Assumption C yields estimated bounds that rule out negative effects of JC training on log wages at week 208. Looking at the IM confidence intervals on the bounds adding Assumption C, we see that with 95 percent confidence the estimated effect is positive. Thus, the effect of JC on log wages at week 208 for EE individuals is statistically positive and between 3.7 and 9.9 percent.

Comparing the second and third columns in Panel B of Table 3, it can be seen that the Full Sample and Non-Hispanics have estimated bounds of similar width, although the bounds for Non-Hispanics are shifted higher to an effect of JC on wages between 5 to 11.8 percent. The IM confidence intervals show that, despite the smaller sample size of Non-Hispanics, this average effect is statistically significant with 95 percent confidence. The estimated bounds for the other demographic groups show some interesting results. All of the bounds and IM confidence intervals exclude zero, with the smallest lower bound being that of the Full Sample at 3.7 percent (all others are 4.6 percent and higher). Remarkably, the estimated bounds for all the demographic groups that exclude Hispanics are relatively similar, suggesting that their average effect of JC on wages for the EE stratum is between 5 and 12 percent. The differences in the confidence intervals across groups is likely driven by the differences in sample sizes. Overall, these results suggest positive average effects of JC on wages across the Non-Hispanic demographic groups, and they reinforce the notion of a strong identifying power of Assumption C.

Given the strong identifying power of Assumption C, it is important to gauge its plausibility. Although a direct statistical test is not feasible, we can indirectly gauge its plausibility by looking at one of its implications. Assumption C formalizes the idea that the EE stratum possesses traits that result in better labor market outcomes relative to individuals in the NE stratum. Thus, we look at pre-treatment covariates that are highly correlated with log wages at week 208 and test whether, on average, individuals in the EE stratum indeed exhibit better characteristics at baseline relative to individuals in the NE stratum. We focus mainly on the following pre-treatment variables: earnings, whether the individual held a job, months employed (all three in the year prior to randomization), and education at randomization.

To implement this idea, we compute average pre-treatment characteristics for the EE and NE strata. Computing average characteristics for the EE stratum is straightforward since, under Assumptions A and B, the individuals in the observed group (T_i, S_i)=(0,1) belong to and are representative of this stratum. Similarly, the mean $E[W|NN]$ can be estimated from the individuals with $(T_i, S_i) = (1, 0)$, who belong to and are representative of the NN stratum. To estimate average characteristics for the NE stratum, note that their average can be written as a function of the averages of the whole population and the other strata, all of which can be estimated under Assumptions A and B. Let W be a pre-treatment characteristic of interest, then,

$$E[W|NE] = \{E[W] - \pi_{EE}E[W|EE] - \pi_{NN}E[W|NN]\}/\pi_{NE}.$$

The estimated differences between the average pre-treatment variables employed for this exercise for the EE and NE strata were all positive, indicating "better" pre-treatment labor market characteristics for the EE stratum. Formal tests of statistical significance for these differences, however, did not reject their equality (mainly because of the high variance in the estimation of $E[W|NE]$). We conclude that this exercise does not provide evidence against Assumption C, while the estimated differences suggest that it is a plausible assumption.[12]

6.3 Narrowing Bounds on ATE_{EE} using a Covariate

We employ earnings in the year prior to randomization as a covariate (X) to narrow the bounds on ATE_{EE} based on the results in Section 3.3. This pre-treatment covariate is

[12]The tables corresponding to this exercise can be found in the Internet Appendix. Employing other pre-treatment variables provided similar results (i.e., no evidence against Assumption C).

highly correlated with log wages at week 208. For each demographic group, we proceeded by splitting the sample into 3 groups based on values of X, each containing roughly the same number of observations. Then, bounds were computed for each group and averaged across groups using the weights $\widehat{\Pr}(X = x_j|EE)$ from Section 3.3.[13]

Table 4 presents the estimated bounds when using earnings in the year prior to randomization to narrow the bounds in Table 3. Panel A shows that the estimated bounds are indeed narrower. The width reduction relative to Panel A of Table 3 is reported in the next to last row. The smallest width reduction is for Blacks at 3.9 percent, while the largest is for Whites at 28 percent. In this application, despite the width reduction, the qualitative results in Panel A of Table 3 are upheld, with the exception of the estimated bounds for Whites, which now exclude zero (although the 95 percent IM confidence interval includes it). Panel B of Table 4 presents narrower estimated bounds under Assumptions A, B and C. Although in this case the procedure only narrows the estimated upper bound that involves trimming, the width reduction achieved (relative to Panel B of Table 3) is comparable to that of the bounds under Assumptions A and B (Panel A of Table 4). As in the top panel, the conclusions that can be gathered from the narrower bounds are qualitatively similar to those in Table 3 (Panel B).

6.4 Bounds on QTE_{EE}^{α} Under Individual-Level Monotonicity

We proceed to analyze the effects of JC on participant's wages beyond the average impact by providing estimated bounds for quantile treatment effects (QTE) for the EE stratum, QTE_{EE}^{α}. Before presenting these bounds, we first briefly discuss HM bounds on the population QTE that are analogous to the HM bounds on the ATE and that employ Assumption A and the empirical bounds of the outcome. To compute HM lower (upper) bounds on the population QTE, we assign to all unemployed individuals in the control group the empirical upper (lower) bound of the outcome and to all unemployed individuals in the treatment group the lower (upper) bound of the outcome. Subsequently, we compare the empirical outcome distributions of both groups by performing quantile regressions of the outcome on the treatment status.[14] For quantiles below 0.45 and above

[13] We decided to employ 3 groups to avoid having too few observations per group in the demographic groups with smaller sample sizes. The results are qualitatively similar when more groups are employed.

[14] Note that the HM bounds on ATE provided at the beginning of Section 6 can be calculated from a linear regression of the outcome on the treatment status after imputing the wages for the unemployed

0.55, the width and value of the estimated bounds on the population QTE (not shown in tables for brevity) are comparable to those of the estimated HM bounds on the ATE presented above. Although we obtain narrower bounds for the quantiles 0.45, 0.50, and 0.55, they remain wide and practically uninformative. For example, the shortest bounds are found for the median and equal [-0.64, 0.64]. The assumptions we consider below substantially tighten these bounds.

To summarize our estimated bounds at several quantiles, we provide a series of figures for the different groups under analysis.[15] The estimated bounds on QTE_{EE}^{α} under Assumptions A and B, along with their corresponding IM confidence intervals, are shown in Figure 1. Recall that the estimated bounds for the ATE_{EE} under the same assumptions presented in Section 6.1 did not rule out zero for any of the groups under analysis. Looking at the estimated bounds on QTE_{EE}^{α} for the Full Sample in Figure 1(a), they rule out zero for all lower quantiles up to 0.7. Once IM confidence intervals are computed, though, only the bounds for the 0.2 quantile imply statistically significant positive effects of JC on log wages with 95 percent confidence. Consistent with the results from bounds on average effects, the estimated bounds on QTE_{EE}^{α} for Non-Hispanics in Figure 1(b) are generally shifted towards the positive space relative to those of the Full Sample. For Non-Hispanics, the estimated bounds also exclude zero for all lower quantiles up to 0.7, and the 95 percent IM confidence intervals rule out zero for the 0.5 quantile. The estimated bounds for these two samples suggest that JC is more likely to have positive effects on log wages for the lower quantiles of the wage distribution.

Looking at the results by race, Figures 1(c) and 1(d) show that the estimated bounds on QTE_{EE}^{α} exclude zero for a number of lower quantiles up to 0.75 (with the exception of the 0.05 quantile for Whites and the 0.75 quantile for Blacks). However, probably due to the smaller sample sizes, when looking at the 95 percent IM confidence intervals for these groups only quantiles 0.55 and 0.65 for Whites and the 0.05 quantile for Blacks are statistically significant. It is worth noting that these two figures suggest that Blacks may experience more positive effects of JC on wages in the lower quantiles of the wage distribution, while Whites may experience more positive effects at the upper quantiles.

Figures 1(e) and 1(f) show the corresponding estimated bounds and 95 percent IM

individuals as described above.

[15] The complete numerical results are shown in the Internet Appendix.

confidence intervals for Non-Hispanic Males and Females, respectively. The bounds reflect a trend of excluding zero at the lower quantiles that is similar to that of the previous groups, albeit less clear for Non-Hispanic Females. Interestingly, Non-Hispanic Males show a greater number of estimated bounds excluding zero, which is probably due to a lower degree of heterogeneity in this group relative to Non-Hispanic Females.[16] Looking at the IM confidence intervals, none of them exclude zero for Non-Hispanic Females, while they do for quantiles 0.05, 0.1, and 0.45 for Non-Hispanic Males. These results suggest that inference for Non-Hispanic Females is more difficult due to their greater heterogeneity and smaller sample size.

To end this subsection, we remark that, while the bounds and IM confidence intervals for the average treatment effect of JC on wages under Assumptions A and B were inconclusive about its sign, the analysis of QTE_{EE}^{α} suggests that positive effects of JC on wages tend to occur for lower and middle quantiles of the distribution. This is the case even when looking at groups with smaller sample sizes. Furthermore, the demographic groups analyzed seem to experience different QTE_{EE}^{α}, both across quantiles and groups. Blacks appear to have larger positive effects at lower quantiles, while Whites appear to have larger effects in the upper quantiles. Also, Non-Hispanic Males show more informative results than Non-Hispanic Females. Next, we add Assumption D (stochastic dominance) to tighten these bounds.

6.5 Bounds on QTE_{EE}^{α} Adding Stochastic Dominance

Estimated bounds on QTE_{EE}^{α} under Assumptions A, B, and D are summarized in Figure 2. The first noteworthy feature of these estimated bounds is that all of them exclude zero at all quantiles, which suggests that the effect of JC on wages is positive along the wage distribution for these groups. These bounds speak to the identifying power of the stochastic dominance assumption (Assumption D). Also noteworthy is that the general conclusions drawn from the estimated bounds in the previous subsection are maintained and reinforced in several instances.

[16]By greater heterogeneity of Non-Hispanic Females relative to Non-Hispanic Males we mean that the former group shows higher standard deviation in key variables such as age, marital and cohabitation status, separated, presence of a child, number of children, and education. This is also true for the average characteristics of the corresponding subset of individuals in the EE stratum.

Looking at the results for the Full Sample and Non-Hispanics (Figures 2(a) and 2(b)), we see again a shift toward more positive effects when Hispanics are excluded. Interestingly, in both of these samples, the lower and upper bounds for the quantiles 0.55 and 0.8 coincide, resulting in a point-identified effect of JC on wages for these two quantiles. Also, adding the stochastic dominance assumption results in 95 percent IM confidence intervals that exclude zero for most of the quantiles except for 0.05, 0.1, 0.6, 0.9, and 0.95 for the Full Sample and 0.1, 0.25, and 0.35 for the Non-Hispanic sample. Concentrating on the latter sample, for which Assumption B is more likely to be satisfied, and excluding the bounds for the quantile 0.05 that differ from the rest, the bounds that exclude zero are between (roughly) 2.7 and 14 percent. In addition, the IM confidence intervals that exclude zero largely overlap, suggesting that the effects of JC on wages do not differ substantially across quantiles. The only clear outliers are the estimated bounds on the 0.05 quantile, which are between 10.5 and 20 percent.

The results by race are shown in Figures 2(c) and 2(d). Adding Assumption D reinforces the notion that Blacks likely exhibit larger positive impacts of JC on log wages in the lower portion of the wage distribution, while Whites likely exhibit larger impacts on the upper quantiles. Indeed, the 95 percent IM confidence intervals for Blacks in the lowest quantiles exclude zero but not those at the highest quantiles. The opposite is true for Whites. However, despite this evidence being stronger than before, it appears inconclusive when looking at the IM confidence intervals, since there is a considerable amount of overlap on the intervals for both groups within quantiles. The IM confidence intervals also show that Blacks have statistically significant positive effects of JC on wages throughout their wage distribution (except at quantiles 0.1, 0.25, 0.9, and 0.95), with estimated bounds that are between 3.1 and 11.5 percent (excluding the 0.05 quantile). Whites show statistically significant positive effects only for quantiles larger than 0.4 (except 0.8), with estimated bounds that are between 6.1 and 14 percent.

Figures 2(e) and 2(f) present the results by Non-Hispanic gender groups. All the estimated bounds for these groups exclude zero at all quantiles, suggesting positive effects of JC on wages and illustrating the identifying power of adding the stochastic dominance assumption. When taking into consideration the 95 percent IM confidence intervals, we find statistically significant positive effects of JC on log wages for more than half of the quantiles considered. Interestingly, Non-Hispanic Females do not have any statistically significant effects throughout the lower half of their wage distribution up to quantile 0.4 (except at the 0.2 quantile), suggesting that Non-Hispanic Females in the upper half of the distribution are more likely to benefit from higher wages due to JC training. Aside from this distinction, there does not seem to be other substantial differences between gender groups, as judged by the large overlap in their IM confidence intervals. Considering confidence intervals that exclude zero, Non-Hispanic Females have estimated bounds that are between 4.4 to 12.1 percent, while those estimated bounds for Non-Hispanic Males are between 3.6 to 13.4 percent (excluding the 0.05 quantile).[17]

6.6 Narrowing Bounds on QTE_{EE}^{α} using a Covariate

To narrow the trimming bounds on QTE_{EE}^{α} we follow the procedure outlined in Section 4.2 employing earnings in the year prior to randomization as a covariate and breaking up the sample into 3 groups, as in Section 6.3. To estimate the bounds on $F_{Y_i(1)|EE}(\widetilde{y})$ in (20), we use 300 values of $\tilde{y}$ that span the support of the outcome. The results are presented in Figures 3 and 4 for bounds under Assumptions A and B, and Assumptions A, B and D, respectively.

The main insights can be summarized as follows. First, reductions in the width of the estimated bounds on QTE_{EE}^{α} are observed in Figures 3 and 4 relative to those in Figures 1 and 2. Although most of the reductions are less than 20 percent, they range from 0 (no reduction) to 100 percent (point identification). Comparing the results in Figures 1 and 3, the reductions in estimated bounds' width across quantiles for the analyzed groups were, on average, 6 percent for the Full Sample, 9 percent for Non-Hispanics, 14 percent for Whites, 2 percent for for Blacks and Non-Hispanic Females, and 16 percent for Non-Hispanic Males. Comparing the results in Figures 2 and 4 that employ stochastic dominance, the reductions in width are more modest as only about a quarter of the estimated bounds' width were reduced (recall that only the lower bounds are subject to trimming and thus to reductions). On average, the reductions in width across quantiles

[17] To indirectly gauge the plausibility of Assumption D in a similar fashion as Assumption C (see Section 6.2), we proceeded to divide each corresponding sample into quintiles based on a given pre-treatment covariate (we employ the same covariates as in Section 6.2). Then, for each quintile we compute and test the difference in the average pre-treatment covariate between the EE and NE strata. As it was the case with Assumption C, we do not find evidence against the stochastic dominance assumption for any of the samples analyzed. The results of this exercise can be found in the Internet Appendix.

were 9 percent for the Full Sample, 14 percent for Non-Hispanics, 10 percent for Whites, 2 percent for Blacks, 4 percent for Non-Hispanic Females, and 11 percent for Non-Hispanic Males. Second, for this empirical application, the reductions in the width of the estimated bounds do not change the qualitative results that were discussed in previous sections. Third, looking at the IM confidence intervals of the estimated bounds that employ X to narrow them, it is evident that in our application the procedure results in wider IM confidence intervals. This is likely due to the required nonparametric estimation of the trimming bounds in (20), which has to be performed for each of the three groups based on X.

6.7 Estimated Bounds for Hispanics

As previously discussed, the original reports in the NJCS found that the program effects on Hispanics' employment and earnings were negative and statistically insignificant (Schochet et al., 2001). This casts doubt on the individual-level monotonicity assumption of the program on employment that was used in analyzing the other demographic groups. The NJCS findings for Hispanics could not be explained by differences (relative to other groups) in baseline characteristics, program participation and degree attainment, duration of enrollment, characteristics of the centers attended, among others (Schochet et al., 2001). Subsequently, Flores-Lagunes et al. (2010) documented that the lack of effect on Hispanics can be partly attributed to the higher local unemployment rates that they face, and to the greater negative impact that Hispanics experience from the local unemployment rates that they face (both factors are especially contrasting relative to Whites).[18] In this section, we do not attempt to provide new explanations for the lack of effect of JC on Hispanics' employment and earnings; instead, we analyze the wage effects of JC for this group by presenting estimated bounds that do not employ the assumption of a non-negative effect of JC on employment (Assumption B).

Zhang and Rubin (2003) and Zhang et al. (2008) show that, under Assumption A (randomly assigned T_i), the four strata proportions in the second column of Table 1 are partially identified as follows (noting that they should sum up to one and are bounded between zero and one): $\pi_{EE} = p_{1|0} - \pi_{EN}$, $\pi_{NE} = p_{1|1} - p_{1|0} + \pi_{EN}$, and $\pi_{NN} = p_{0|1} - \pi_{EN}$,

[18] Additionally, Flores-Lagunes et al. (2010) document that JC appears to "shield" Whites from the effects of adverse local unemployment rates, but not Hispanics or Blacks.

with π_{EN} satisfying $max(0, p_{1|0} - p_{1|1}) \leq \pi_{EN} \leq min(p_{1|0}, p_{0|1})$. These bounds on the strata proportions can be used to construct bounds on ATE_{EE} and QTE^{α}_{EE} that do not require Assumption B; however, the resulting bounds are expected to be wide given the considerable identifying power of Assumption B.[19] To conserve space, the expressions of the bounds that do not use Assumption B are relegated to the Internet Appendix.

Table 5 reports, for the Hispanic sample, two sets of estimated bounds on the ATE_{EE} of JC on log wages that do not employ Assumption B. The second column shows estimated bounds under Assumption A only, while the third column adds Assumption C. As expected, the estimated bounds are wide: -0.451 to 0.359 (under Assumption A) and -0.448 to 0.359 (under Assumptions A and C). Interestingly, adding Assumption C now results in a fairly small tightening of the bounds. This is in contrast to the results presented in the previous sections that made use of Assumption B, where Assumption C had considerable identifying power. Figure 5 reports the estimated bounds on QTE^{α}_{EE} for Hispanics under Assumption A (top panel) and under Assumptions A and D (bottom panel). Three points are noteworthy. First, as it was the case with the bounds on ATE_{EE}, the bounds without Assumption B are wide across the quantiles analyzed. Second, adding Assumption D now results in a very small tightening of the bounds, as it was the case above with Assumption C. Third, despite the wideness of both sets of estimated bounds for Hispanics, a pattern emerges in which they are considerably narrower for the upper part of the distribution of log wages. Thus, while positive or negative effects of JC on log wages for Hispanics cannot be ruled out, both large (but potentially plausible) negative and positive effects can be ruled out for the upper part of their wage distribution.

Finally, for comparison purposes, we computed bounds that do not employ Assumption B for all other demographic groups (which are available in the Internet Appendix). These estimated bounds (for both ATE_{EE} and QTE^{α}_{EE}) are also wide and include zero, although they are less wide than those for Hispanics. Hence, the inconclusiveness of the estimated effects of JC on log wages for Hispanics can be due to a true lack of effect for them or to the relative uninformativeness of the assumptions being employed to estimate their bounds.

[19] Recall that Assumption B allows point identification of not only the strata proportions to trim the data, but also of $E[Y_i(0)|EE]$ in (6) and $F^{-1}_{Y_i(0)|EE}(\alpha)$ in (14).

7 Conclusion

We review and extend recent nonparametric bounds for average and quantile treatment effects that account for sample selection, and that require weaker assumptions than those conventionally employed for point identification of these effects. These techniques are applied to the problem of assessing the effect of Job Corps (JC) training on wages accounting for non-random selection into employment. Since JC's stated goal is to enhance participants' human capital and labor market outcomes, research shedding light on the effects of JC on wages is important because positive wage effects can be related to human capital improvements. Under the assumptions we consider, our results suggest that JC has positive and statistically significant effects on wages for the individuals who would be employed regardless of participation in JC, not only at the mean but also at different points of the wage distribution, and for different demographic groups of interest (with the exception of Hispanics).

We start by exploiting the random assignment into the program in our data to construct Horowitz and Manski (2000) bounds, and then add an individual-level monotonicity assumption on the effect of JC on employment to tighten them. While the latter bounds cannot rule out negative average effects of JC on wages for those employed irrespective of treatment assignment, by constructing bounds on quantile treatment effects we find that for certain quantiles and demographic groups we are able to statistically rule out zero or negative effects. These results are noteworthy given that the lower bound under these assumptions is likely too pessimistic since it implies a theoretically implausible perfect negative correlation between wages and employment.

To further tighten the above bounds, we add assumptions formalizing the notion that individuals in some strata are likely to have better labor market outcomes than others, hence avoiding the perfect negative correlation between wages and employment implied by the previous bounds. The estimated bounds for the average effect of JC on wages for the individuals employed irrespective of treatment assignment suggest statistically significant positive effects. The estimated bounds for groups that exclude Hispanics are remarkably similar, with an estimated lower bound of about 4.6 percent and an upper bound of about 12 percent. We obtain interesting insights when analyzing bounds on quantile treatment effects for individuals employed irrespective of treatment assignment. In particular, our results suggest that the positive effects of JC on wages largely hold across quantiles, but that there are differences across quantiles and demographic groups. The effects for Blacks appear larger in the lower half of their wage distribution, while the effects appear larger for Whites in the upper half of their wage distribution. In addition, Non-Hispanic Females show statistically significant positive effects of JC on wages in the upper part of their wage distribution, but not in the lower part. Our preferred estimated bounds on quantile effects—those imposing individual-level monotonicity and stochastic dominance—for the Non-Hispanic groups suggest that the statistically significant effects of JC on wages across quantiles range from about 2.7 to 14 percent. We also discuss how to employ a pre-treatment covariate to narrow these bounds, which in our application results in average width reductions of about 10 percent for ATE_{EE} and 8 percent for QTE^{α}_{EE}.

For Hispanics we conduct a separate analysis that does not employ the assumption of individual-level monotonicity of the effect of JC on employment, since prior evidence suggests that this assumption is less likely to hold for them. Estimated bounds without this assumption are wide and include zero, implying that we are unable to rule out zero, negative, or positive effects of JC on wages for Hispanics. However, estimated bounds across quantiles of the wage distribution indicate that large (but potentially plausible) positive and negative effects can be ruled out for the upper quantiles of this group.

In general, our application illustrates the usefulness of the nonparametric bounds discussed in this paper in settings where sample selection is present, as well as the insights that can be gained from employing these techniques to analyze quantile treatment effects. For instance, consider the example given in the Introduction about the antismoking assistance program for pregnant women studied in Sexton and Hebel (1984). In this case, the present methods could be used to bound the average and quantile effects of the program on the birth weight of those babies who would not die during gestation regardless of their mothers' participation in the program. In addition, the JC application points toward some caveats of the approach and important directions in which these methods could be extended. Although similar assumptions are commonly used in the literature, the (untestable) individual-level monotonicity assumption of the treatment on the selection indicator is non-trivial and can be hard to justify in practice. Similarly, the bounds can be affected by measurement error. Therefore, the derivation of tighter bounds that do not rely on that monotonicity assumption and an analysis of the effects of measurement error on the bounds we consider would be valuable econometric contributions to the literature.

8 References

Abadie, A., Angrist, J., and Imbens, G. 2002. "Instrumental Variables Estimation of Quantile Treatment Effects." Econometrica, 70: 91-117.

Angrist, J., Chen, S., and Frandsen, B. 2009. "Did Vietnam Veterans Get Sicker in the 1990s? The Complicated Effects of Military Service on Self-Reported Health." National Bureau of Economic Research Working Paper No. 14781.

Angrist, J., and Krueger, A. 1999. "Empirical Strategies in Labor Economics." In Orley Ashenfelter and David Card (eds) Handbook of Labor Economics, Volume IIIA, Elsevier.

Angrist, J., and Krueger, A. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." Journal of Economic Perspectives, 15: 69-85.

Bhattacharya, J., Shaikh, A., and Vytlacil, E. 2008. "Treatment Effects Bounds Under Monotonicity Assumptions: An Application to Swan-Ganz Catheterization." American Economic Review: Papers and Proceedings, 98: 351-356.

Blundell, R., Gosling, A., Ichimura, I., and Meghir, C. 2007. "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds." Econometrica, 75: 323-363.

Bollinger, C. 1996. "Bounding Mean Regressions when a Binary Regressor is Mismeasured." Journal of Econometrics, 73: 387-399.

Dobkin, C., and Shabani, R. 2009. "The Health Effects of Military Service: Evidence from the Vietnam Draft". Economic Inquiry, 47: 69-80.

Eisenberg, D., and Rowe, B. 2009. "The Effect of Smoking in Young Adulthood on Smoking Later in Life: Evidence based on the Vietnam Era Draft Lottery". Forum for Health Economics & Policy, 12: Article 4.

Flores, C., and Flores-Lagunes, A. 2010. "Nonparametric Partial Identification of Causal Net and Mechanism Average Treatment Effects.", Mimeo, University of Miami.

Flores, C., Flores-Lagunes, A., Gonzales, A., and Neumann, T. 2012. "Estimating the Effects of Length of Exposure to Instruction in a Training Program: The Case of Job

Corps." The Review of Economics and Statistics, 94: 153-171.

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.

Frangakis, C., and Rubin, D. 2002. "Principal Stratification in Causal Inference." Biometrics, 58: 21-29.

Gundersen, C., Kreider, B., and Pepper, J. 1995. "The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis." Journal of Econometrics, 166: 79-91.

Hearst, N., Newman, T., and Hulley, S. 1986. "Delayed Effects of the Military Draft on Mortality; A Randomized Natural Experiment." The New England Journal of Medicine, 314: 620-624.

Heckman, J. 1979. "Sample Selection Bias as a Specification Error." Econometrica, 47: 153-162.

Heckman, J. 1990. "Varieties of Selection Bias." American Economic Review, 80: 313-318.

Heckman, J., LaLonde, R., and Smith, J. 1999. "The Economics and Econometrics of Active Labor Market Programs." In Orley Ashenfelter and David Card (eds.) Handbook of Labor Economics, Volume IIIA, Elsevier.

Heckman, J., and Smith, J. A. 1995. "Assessing the Case for Social Experiments." Journal of Economic Perspectives, 9(2): 85-110.

Heckman, J., and Smith, J. A. 1999. "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme: Implications for Simple Programme Evaluation Strategies." Economic Journal, 109(2): 313-348.

Horowitz, J., and Manski, C. 1995. "Identification and Robustness with Contaminated and Corrupted Data." Econometrica, 63: 281-302.

Horowitz, J., and Manski, C. 2000. "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data." Journal of the American Statistical Association, 95: 77-84.

Imai, K. 2008. "Sharp Bounds on the Causal Effects in Randomized Experiments with "Truncation-by-Death"." Statistics and Probability Letters, 78: 144-149.

Imbens, G., and Angrist, J. 1994. "Identification and Estimation of Local Average Treatment Effects." Econometrica, 62: 467-476.

Imbens, G., and Manski, C. 2004. "Confidence Intervals for Partially Identified Parameters." Econometrica, 72: 1845-1857.

Imbens, G., and Wooldridge, J. 2009. "Recent Developments in the Econometrics of Program Evaluation." Journal of Economic Literature, 47: 5-86.

Lechner, M., and Melly, B. 2010. "Partial Identification of Wage Effects of Training Programs." Mimeo, University of St. Gallen.

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

Manski, C. 1994. "The Selection Problem." in C. Sims (ed) Advances in Econometrics, Sixth World Congress, vol I, Cambridge, U.K. Cambridge University Press, 143-170.

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

Molinari, F. 2008. "Partial Identification of Probability Distributions with Misclassified Data." Journal of Econometrics, 144: 81-117.

Schochet, P. 2001. "National Job Corps Study: Methodological Appendixes on the Impact Analysis." Mathematica Policy Research, Inc., Princeton, NJ.

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.

Sexton, M., and Hebel, R. 1984. "A Clinical Trial of Change in Maternal Smoking and Its Effect on Birth Weight." Journal of the American Medical Association, 251: 911-915.

US Department of Labor. 2010. http://www.dol.gov/dol/topic/training/jobcorps.html.

USA Today. 2011. "Training Sprawl Costs U.S. $18 Billion per Year", February 9, 2011.

van Ours, J. 2004. "The Locking-in Effect of Subsidized Jobs." Journal of Comparative Economics, 32: 37-52.

Zhang, J., and Rubin, D. 2003. "Estimation of Causal Effects via Principal Stratification When Some Outcomes are Truncated by 'Death'." Journal of Educational and Behavioral Statistics, 28: 353-368.

Zhang, J., Rubin, D., and Mealli, F. 2008. "Evaluating the Effect of Job Training Programs on Wages Through Principal Stratification." in D. Millimet et al. (eds) Advances in Econometrics vol XXI, Elsevier.

Zhang, J., Rubin, D., and Mealli, F. 2009. "Likelihood-based Analysis of the Causal Effects of Job Training Programs Using Principal Stratification." Journal of the American Statistical Association, 104: 166-176.

Table 1. Observed groups based on treatment and employment indicators (T_i, S_i) and Principal Strata (PS) mixture within groups.

Observed groups by (T_i, S_i)	Principal Strata (PS)	PS after imposing Individual-level monotonicity
(0,0)	*NN* and *NE*	*NN* and *NE*
(1,1)	*EE* and *NE*	*EE* and *NE*
(1,0)	*NN* and *EN*	*NN*
(0,1)	*EE* and *EN*	*EE*

Table 2. Estimated principal strata proportions by demographic groups under analysis.

PS	Full Sample	Non-Hispanics	Whites	Blacks	Non-Hispanic Males	Non-Hispanic Females	Hispanics
EE	0.566	0.559	0.657	0.512	0.583	0.530	0.598
NN	0.393	0.392	0.303	0.436	0.377	0.410	0.400
NE	0.041	0.049	0.040	0.052	0.040	0.060	0.002
Observations	9145	7573	2358	4566	4280	3293	1572

Note: All estimated proportions are statistically significant at the 1 percent level, except the proportion of NE individuals for the group of Hispanics.

Table 3. Bounds on the average treatment effect of the *EE* stratum for log wages at week 208, by demographic groups.

Panel A: Under Assumptions A and B

	Full Sample	Non-Hispanics	Whites	Blacks	Non-Hispanic Females	Non-Hispanic Males
Upper bound	0.099	0.118	0.120	0.116	0.120	0.114
	(0.014)	(0.015)	(0.028)	(0.020)	(0.024)	(0.020)
Lower bound	-0.022	-0.018	0.000	-0.012	-0.023	-0.009
	(0.016)	(0.017)	(0.031)	(0.021)	(0.026)	(0.023)
Width	0.121	0.136	0.120	0.129	0.143	0.123
95 percent IM confidence interval	[-0.049, 0.122]	[-0.046, 0.143]	[-0.050, 0.166]	[-0.047, 0.149]	[-0.066, 0.159]	[-0.047, 0.147]

Panel B: Under Assumptions A, B, and C

	Full Sample	Non- Hispanics	Whites	Blacks	Non-Hispanic Females	Non-Hispanic Males
Upper bound	0.099	0.118	0.120	0.116	0.120	0.114
	(0.014)	(0.015)	(0.028)	(0.020)	(0.024)	(0.020)
Lower bound	0.037	0.050	0.056	0.053	0.046	0.052
	(0.012)	(0.013)	(0.022)	(0.016)	(0.020)	(0.016)
Width	0.062	0.068	0.064	0.063	0.074	0.061
95 percent IM confidence interval	[0.018, 0.122]	[0.029, 0.143]	[0.019, 0.166]	[0.027, 0.149]	[0.014, 0.159]	[0.026, 0.147]

Note: Bootstrap standard errors in parentheses (based on 5,000 replications). IM refers to the Imbens and Manski (2004) confidence interval, which contains the true value of the parameter with a given probability.

Table 4. Bounds on average treatment effect of the EE stratum for log wages at week 208, by demographic groups, employing earnings in the year prior to randomization (*X*) to narrow the bounds.

Panel A: Under Assumptions A and B

	Full Sample	Non-Hispanics	Whites	Blacks	Non-Hispanic Females	Non-Hispanic Males
Upper bound	0.096	0.113	0.102	0.113	0.117	0.107
	(0.014)	(0.015)	(0.025)	(0.019)	(0.023)	(0.019)
Lower bound	-0.018	-0.012	0.015	-0.010	-0.018	-0.002
	(0.016)	(0.017)	(0.027)	(0.020)	(0.025)	(0.021)
Width	0.114	0.125	0.087	0.123	0.134	0.109
Width reduction (vs. Table 3.A)	5.7%	8.0%	27.8%	3.9%	5.9%	11.4%
95 percent level IM confidence interval	[-0.043, 0.119]	[-0.040, 0.138]	[-0.030, 0.144]	[-0.043, 0.144]	[-0.059, 0.155]	[-0.038, 0.138]

Panel B: Under Assumptions A, B, and C

	Full Sample	Non- Hispanics	Whites	Blacks	Non-Hispanic Females	Non-Hispanic Males
Upper bound	0.096	0.113	0.102	0.113	0.117	0.107
	(0.014)	(0.015)	(0.025)	(0.019)	(0.023)	(0.019)
Lower bound	0.037	0.050	0.056	0.053	0.046	0.052
	(0.012)	(0.013)	(0.022)	(0.016)	(0.020)	(0.016)
Width	0.059	0.063	0.046	0.060	0.071	0.054
Width reduction (vs. Table 3.B)	4.9%	7.3%	28.2%	4.8%	4.3%	11.4%
95 percent level IM confidence interval	[0.018, 0.119]	[0.029, 0.138]	[0.019, 0.144]	[0.027, 0.144]	[0.014, 0.155]	[0.026, 0.138]

Note: Bootstrap standard errors in parentheses (based on 5,000 replications). IM refers to the Imbens and Manski (2004) confidence interval, which contains the true value of the parameter with a given probability.

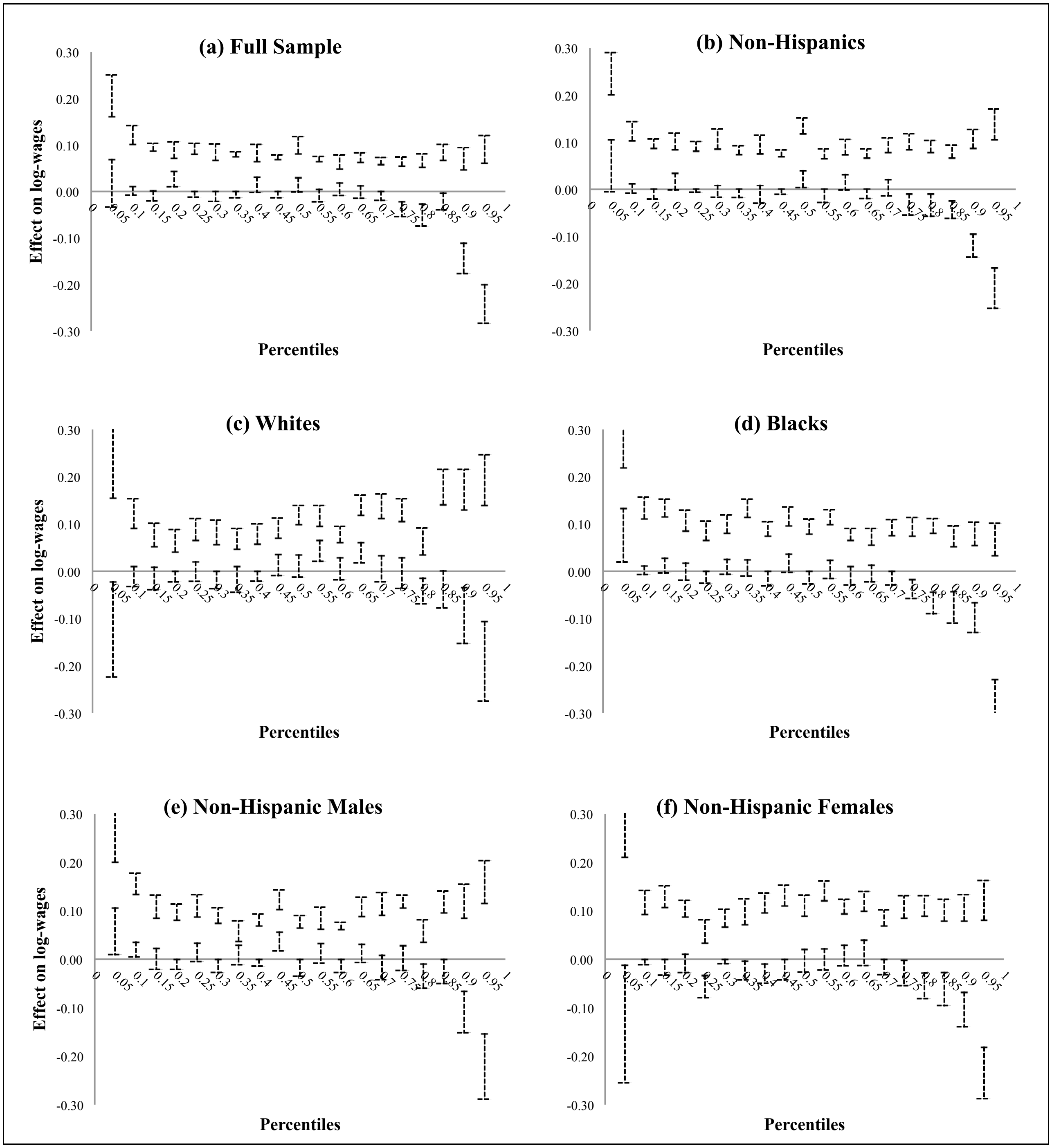

Figure 1. Bounds and 95 percent Imbens and Manski (2004) confidence intervals for *QTE* of the EE stratum by demographic groups, under Assumptions A and B. Upper and lower bounds are denoted by a short dash, while IM confidence intervals are denoted by a long dash at the end of the dashed vertical lines.

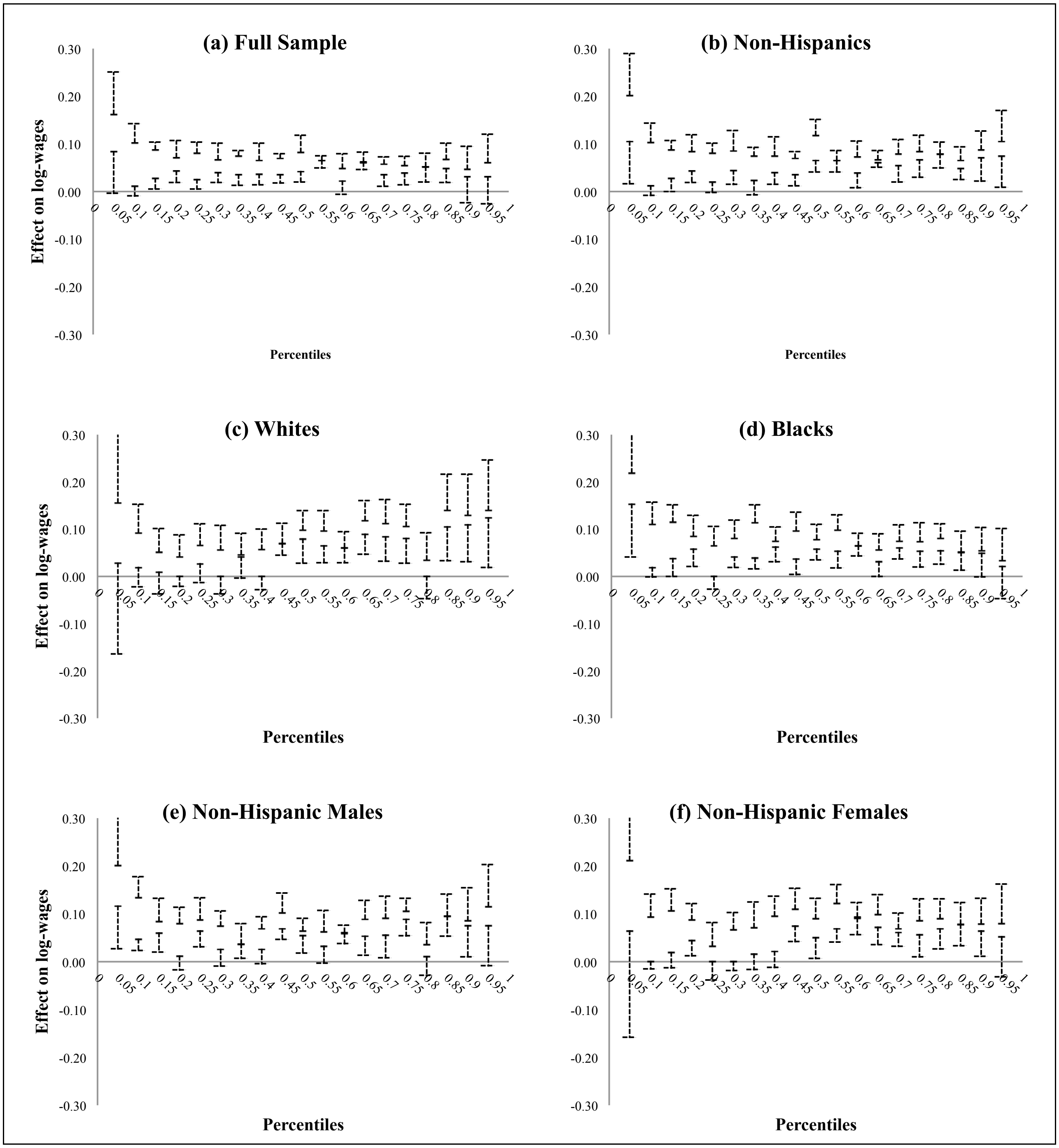

Figure 2. Bounds and 95 percent Imbens and Manski (2004) confidence intervals for *QTE* of the EE stratum by demographic groups, under Assumptions A, B, and D. Upper and lower bounds are denoted by a short dash, while IM confidence intervals are denoted by a long dash at the end of the dashed vertical lines.

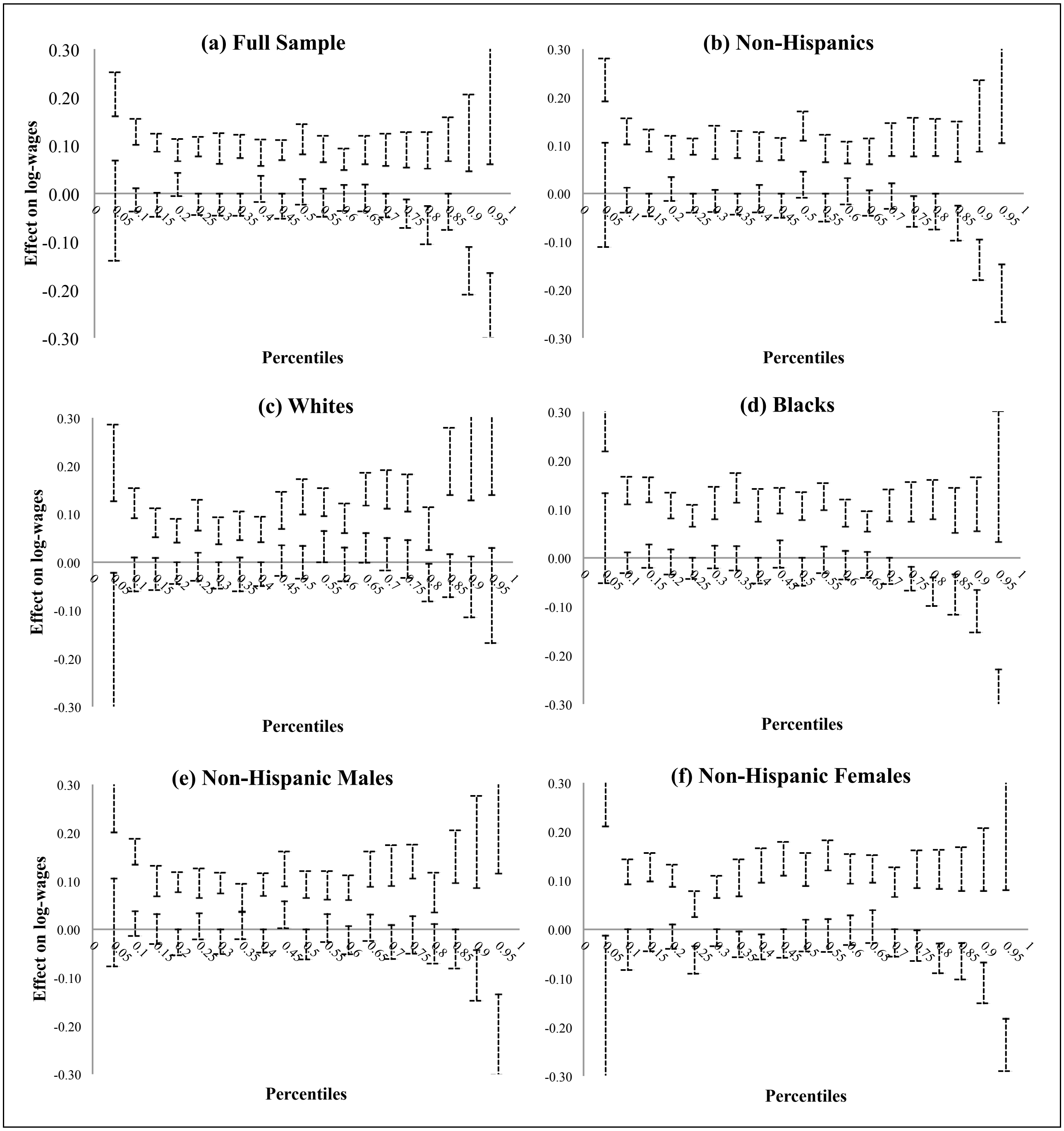

Figure 3. Bounds and 95 percent Imbens and Manski (2004) confidence intervals for *QTE* of the EE stratum by demographic groups, under Assumptions A and B, using earnings in the year prior to randomization as a covariate to narrow the bounds. Upper and lower bounds are denoted by a short dash, while IM confidence intervals are denoted by a long dash at the end of the dashed vertical lines.

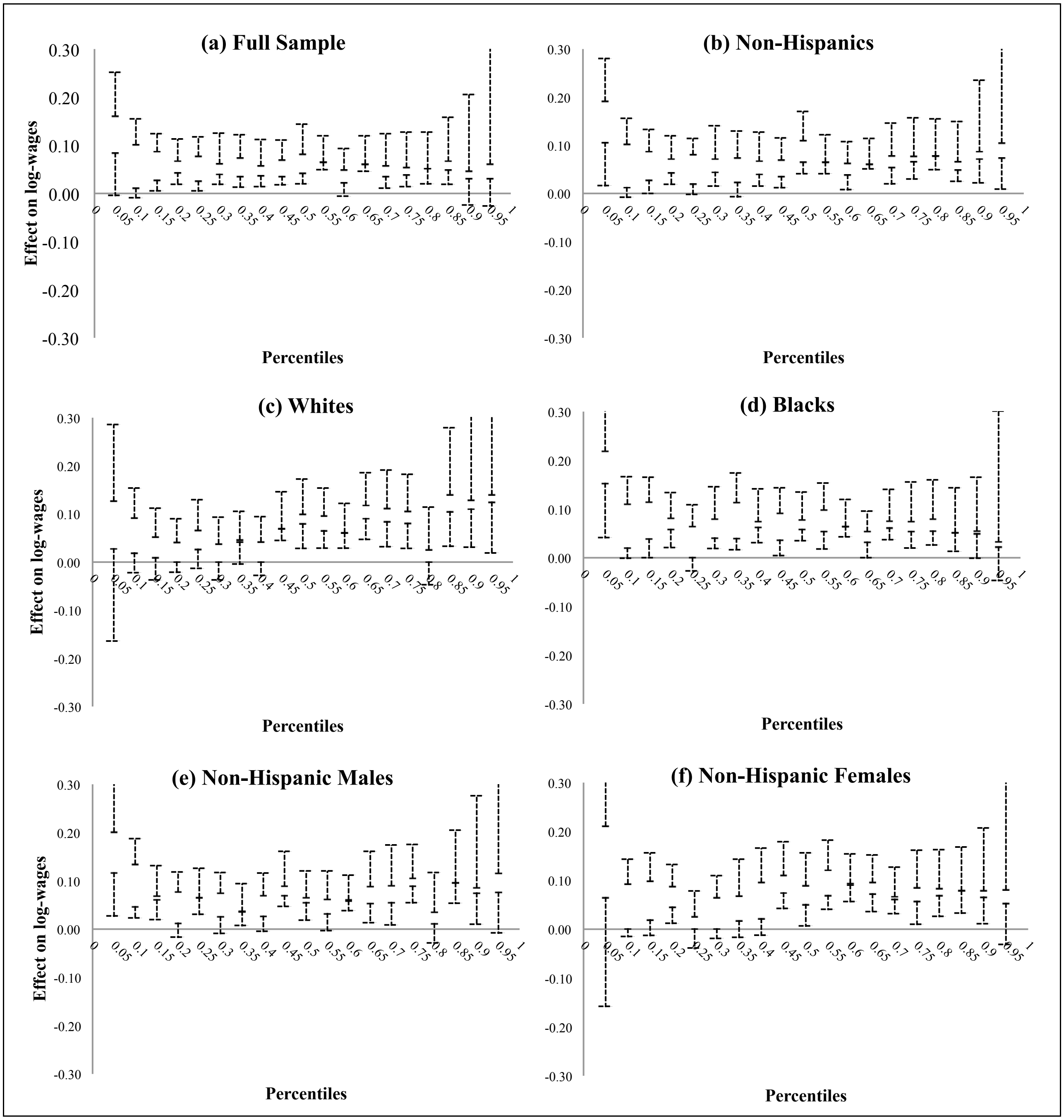

Figure 4. Bounds and 95 percent Imbens and Manski (2004) confidence intervals for *QTE* of the EE stratum by demographic groups, under Assumptions A, B and D, using earnings in the year prior to randomization as a covariate to narrow the bounds. Upper and lower bounds are denoted by a short dash, while IM confidence intervals are denoted by a long dash at the end of the dashed vertical line.

Table 5. Bounds on the average treatment effect of the *EE* stratum for log wages at week 208 for Hispanics.

	Assumption A	Assumption A & C
Upper Bound	0.359	0.359
	(0.044)	(0.044)
Lower Bound	-0.451	-0.448
	(0.053)	(0.049)
Width	0.810	0.807
95 percent level IM confidence interval	[-0.538, 0.431]	[-0.528, 0.431]

Note: Bootstrap standard errors in parentheses (with 5,000 replications)

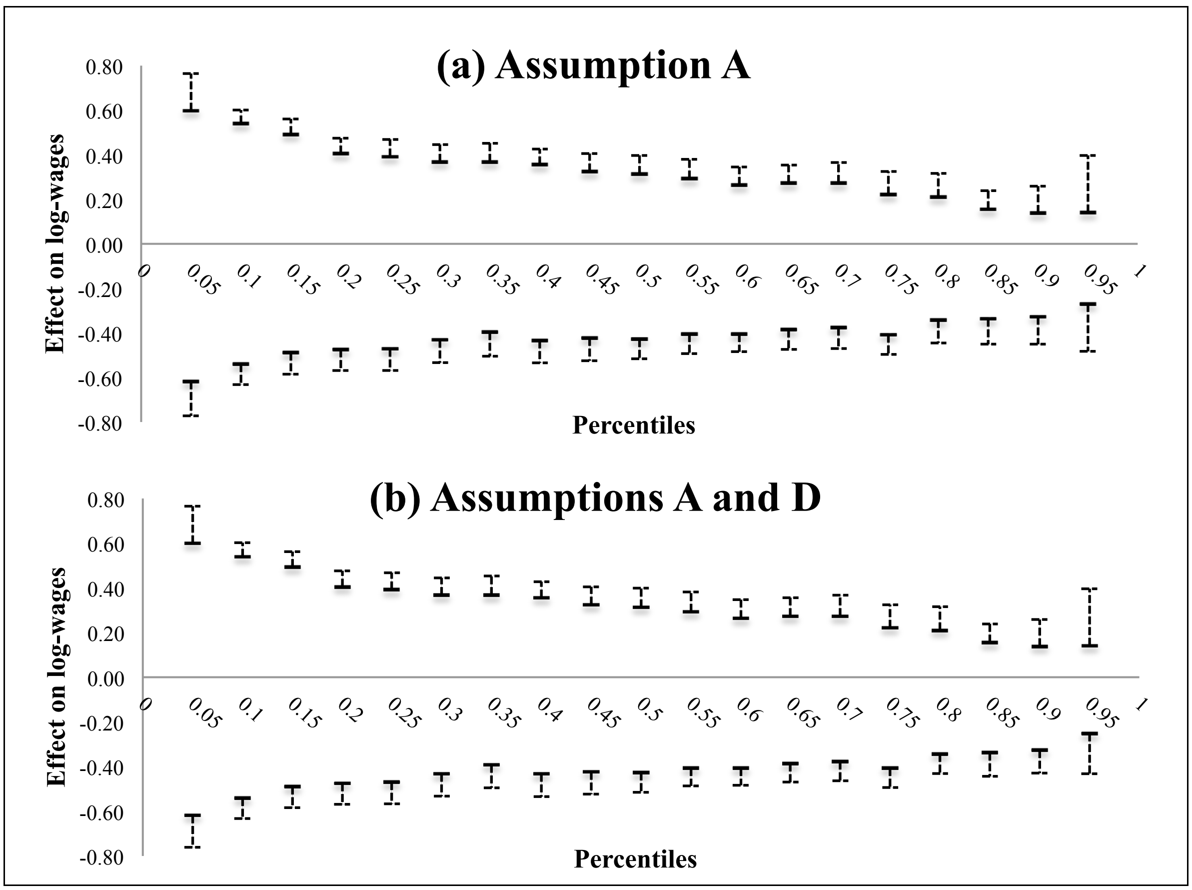

Figure 5. Bounds and 95 percent Imbens and Manski (2004) confidence intervals for *QTE* of the EE stratum for Hispanics, under (a) Assumption A, and (b) Assumptions A and D. Upper and lower bounds are denoted by a short dash, while IM confidence intervals are denoted by a long dash at the end of the dashed vertical lines.