Bounding Average and Quantile Effects of Training on Employment and Unemployment Durations under Selection, Censoring, and Noncompliance

German Blanco[*], Xuan Chen[†], Carlos A. Flores[‡] and Alfonso Flores-Lagunes[§]

Preliminary Draft, not for circulation.

April 14, 2017

Abstract

Using data from a randomized evaluation of the Job Corps (JC) training program, we estimate nonparametric bounds for average and quantile treatment effects of training on employment and unemployment duration. Under relatively weak assumptions, we bound these effects addressing three pervasive problems in randomized evaluations: sample selection, censoring, and noncompliance. The first arises when the individuals' decision to experience employment or unemployment spells is endogenous and potentially affected by the program. Censoring arises when the duration outcome is fully observed only for individuals who have completed a full spell by the end of the observation period, with the extent of censoring being potentially affected by training. Noncompliance is present when some assigned participants do not receive training and some assigned nonparticipants receive training. Ignoring these issues would yield biased estimates of the effects. Our results indicate that JC training increases the average duration in weeks of the last complete employment spell before week 208 after randomization by at least 10.7 log points (11.3 percent) for individuals who comply with their treatment assignment and who would experience a complete employment spell whether or not they enrolled in JC. The proposed approach allows us to also bound the wage effects of JC for these individuals during those spells. We find that JC increases their average wages by between 6.2 and 13.7 log points (6.4 and 14.7 percent), suggesting that JC not only helps these individuals to maintain their jobs longer, but also that those jobs are better paid. We find no distinguishable effects of JC on average unemployment duration. The quantile results reflect heterogeneous effects and strengthen our conclusions based on the average effects.

[*]Department of Economics, Illinois State University (gblanco@ilstu.edu).

[†]School of Labor and Human Resources, Renmin University of China (xchen11@ruc.edu.cn).

[‡]Department of Economics, California Polytechnic State University at San Luis Obispo (cflore32@calpoly.edu).

[§]Senior Research Associate, CPR; Economics Department, Syracuse University; and IZA Research Fellow (afloresl@maxwell.syr.edu).

1 Introduction

From an economic and policy making perspective, evaluating the effectiveness of active labor market programs is of extreme interest. The program evaluation literature is concerned with such an important task. The methodological and empirical work found in the forgoing literature vastly focuses on estimating average effects of a policy variable on non-duration labor market outcomes such as earnings and employment (for a review see Heckman et al., 1999; and Imbens and Wooldridge, 2009). While estimating the effects of labor market programs on employment rates is important, it is also crucial from a policy point of view to analyze their effects on the duration of employment and unemployment spells. As discussed by Ham and LaLonde (1996), such programs may improve employment rates by helping unemployed individuals find jobs faster (e.g., by improving their job search skills) or by helping employed individuals keep their jobs longer (e.g., by improving their wok habits). Hence, estimating program effects on the duration of employment and unemployment spells can shed light into how the program affects employment rates, providing valuable information for policy makers (Ham and LaLonde, 1996). Estimating such effects, however, is challenging. Even when employing data from an experimental evaluation, assessing the program impact on duration outcomes is often complicated by three identification problems: sample selection, censoring, and noncompliance. The first problem arises because, even when individuals are randomized at baseline, who experiences an employment or unemployment spell post-randomization is not random and is potentially affected by training participation. Censoring occurs because the full duration of employment and unemployment spells is fully observed only if these spells are completed before the end of the observation period, with the extent of censoring also being potentially affected by training participation. Finally, noncompliance occurs because some individuals assigned to participate in the program do not enroll in it, while some controls individuals do enroll. In this paper, we bound the average and quantile effects of an important training program for disadvantaged youth in the U.S.—Job Corps—on the duration of its participants' employment and unemployment spells in the presence of sample selection, censoring, and noncompliance.

The Job Corps (JC) program is America's largest and most comprehensive education and job training program enrolling disadvantaged youth, ages 16 to 24, at no cost to them. It offers a myriad of services such as academic, vocational, and skills training, health care, health education, counseling, and job search assistant at its more than 120 centers nationwide—where most of its participants reside during training. Federal funds to keep the program running are around \$1.6 billion per year (US Department of Labor, DOL, 2013), which makes its evaluation of paramount public interest. To determine the program's effectiveness, during 1995-96 the DOL funded the National Job Corps Study (NJCS), whose main feature was the random assignment of eligible participants to a treatment or control group.[1] Previous research using the NJCS has found positive effects of JC on employment rates (e.g., Schochet et al., 2001; Lee, 2009; Frumento et al., 2012; Blanco et al., 2013a; Chen and Flores, 2015; Chen et al., 2017). However, despite the importance of analyzing the effects of training programs on the duration of employment and unemployment spells, as previously discussed, to the best of our knowledge no previous study has assessed the effects of this major program on such outcomes. This paper fills this void.

Using data from the NJCS, various studies have examined the average effectiveness of being assigned to enroll in JC (rather than actual enrollment in JC), where most of the studied labor market outcomes were non-duration outcomes such as earnings, employment, and wages (e.g., Schochet et al., 2001; Lee, 2009; Zhang et al., 2009; Flores-Lagunes et al., 2010; Flores and Flores-Lagunes, 2010, 2013; Blanco et al., 2013a, 2013b). This parameter, typically referred to as the Intention to Treat Effect (ITT), does not capture the actual effect of receiving training when there is noncompliance in treatment/control assignment. This was the case in the NJCS, where only about 73.8 percent of the individuals assigned to the treatment group participated in JC, while about 4.4 percent of the individuals assigned to the control group participated in JC. A relevant parameter that explicitly considers noncompliance is the Local Average Treatment Effect or LATE (Imbens and Angrist, 1994; Angrist et al., 1996), which estimates the effect of participating in the program by using the treatment assignment indicator as an instrumental variable for actual participation. Studies analyzing effects of actual JC participation on earnings , employment, and wages include, among others, Schochet et al. (2001), Frumento et al. (2012), Eren and Ozbeklik (2014), and Chen and Flores (2015).[2] Here, we analyze the effects of JC on employment and unemployment duration spells, which in addition to the noncompliance problem also face sample selection and censoring.[3] In addition, unlike most of these papers (with the exception of Eren and Ozbeklik, 2014), we go beyond the estimation of average effects and also consider quantile effects; that is, we also analyze the effects of JC at different points of the distribution of the duration of employment and unemployment spells.

[1] It is well-known that the use of experimental data facilitates the identification of causal effects under certain conditions. Non-experimental techniques have also been developed within the program evaluation literature. There is, however, debate about the reliability of these techniques (e.g., LaLonde, 1986; Dehejia and Wahba 1999, 2002; Smith and Todd, 2005; Dehejia, 2005), which makes experimental evaluations even more appealing.

[2] Eren and Ozbeklik (2014) focus on estimating average and quantile local effects of JC participation on earnings 48 months after randomization controlling for noncompliance by employing the methods in Frölich and Melly (2013).

[3] As discussed below, Frumento et al. (2012) and Chen and Flores (2015) also address the sample selection problem when assessing the effect of JC participation on wages.

A difficulty of analyzing duration outcomes like the ones analyzed here stems from the fact that some of the spells are not fully observed, i.e., they are censored at the end of the observation period. Even when employing experimental data, as in our application, estimating the impacts of treatment assignment and actual training is not straightforward because the extent of censoring is potentially affected by actual participation in the program. Ignoring censoring will result in biased estimates. One conventional approach to address censoring (and selection) is based on the proportional hazard model, which usually relies heavily on parametric assumptions about the functional form of the relations of interest (e.g., Eberwein et al., 1997, 2002; Ham and LaLonde, 1996; Heckman and Singer, 1984, whose approach was semiparametric; Abbring and Van den Berg, 2003, who proposed a nonparametric version of the mixed proportional hazard model while imposing a multiplicative structure on the hazard rate to model unobserved heterogeneity). Another literature follows and extends the work of Powell (1986) and Chernozhukov and Hong (2002), among others, on censored quantile regression (CQR). For example, work on CQR has been extended by Blundell and Powell (2007), Chernozhukov et al. (2015), and Frandsen (2015) to deal with an endogenous regressor by employing instrumental variables based estimators. Whether the method is parametric, semi or nonparametric, assuming independence of censoring points is standard in models based on CQR.[4] As stated before, in this application we allow for the extent of censoring to be affected by actual participation in the program.

In this paper, we take an alternative nonparametric approach that relies on relatively weak assumptions to construct bounds on the average and quantile treament effects of actual training receipt on the duration of employment and unemployment spells within a principal stratification framework (Frangakis and Rubin, 2002) to address the three identification issues previously discussed: selection, censoring, and noncompliance. Principal stratification—which has its roots in the instrumental variables analysis by Angrist et al. (1996)—provides a framework for analyzing treatment effects when controlling for variables that have been affected by the treatment (e.g., selection into employment spells or censoring). It is based on the idea of comparing individuals who share the same potential values of the post-treatment variable(s) one wants to adjust for (e.g., selection into employment spells) under both treatment arms. Previously, bounds have been used to analyze the effects of training programs within the spirit of this framework. Zhang et al. (2008) and Lee (2009) corrected for sample selection in their analysis of average wages (as wages are observed only for employed individuals), but did not account for noncompliance

[4]An alternative two-step estimator based on Kaplan-Meier integrals is proposed in Sant'anna (2016), who also employs the standard condition about independent censoring points found in the CQR literature.

and censoring was not present. Lee (2009) employed his bounds to estimate the wage effects of JC. Blanco et al. (2013a, 2013b) used the same bounds as in Zhang et al. (2008) and Lee (2009), and their extension to quantiles by Imai (2008) to analyze the wage effects of JC, however, they did not extend their analysis to explicitly account for noncompliance. Focusing on average effects, Chen and Flores (2015) derived bounds accounting for both noncompliance and sample selection, and employed them to analyze the wage effects of JC. The general finding in those papers assessing the average wage effects of JC is that this effect is positive four years after randomization. Importantly, none of those papers considered duration outcomes. There is, however, another strand of the literature that bounds duration outcomes. For example, using a different framework and set of assumptions, Vikström et al. (2016) propose bounds on transition probabilities, although they do not allow for noncompliance. In contrast, our focus is on the average and quantile effects of actual training on time to transitions (i.e., employment and unemployment durations), rather than transition probabilities.[5]

The approach proposed in this paper to analyze the average and quantile effects of training on employment and unemployment durations is based on the sharp bounds derived by Imai (2007) to address noncompliance and selection, which, by the way we define our parameters of interest in our setting, also allow us to account for censoring. These bounds are based on monotonicity and stochastic dominance assumptions, and assume there is a valid instrument to address noncompliance (but not to address sample selection). While the bounds employed herein are based on those in Imai (2007), most of them differ from his because we consider different subpopulations and the direction of our monotonicity and stochastic dominance assumptions also differs. Thus, the bounds presented herein complement those in Imai (2007).

Our results suggest that JC training has a positive and statistically significant average effect on the duration in weeks of the last complete employment spell before week 208 after randomization for individuals who comply with their treatment assignment (the "compliers") and who would experience a complete employment spell whether or not they enrolled in JC, a subpopulation representing about 26 percent of our main sample—which excludes Hispanics—and 37 percent of all compliers. Under our preferred set of assumptions, this effect is bounded between 10.7 and 46.5 log points (11.3 and 59.2 percent). In addition, we note that the effect is heterogeneous along the employment duration distribution. In some of the lower percentiles the effect is positive and statistically significant; in contrast, while positive effects can potentially happen

[5]In addition to employing random assignment of treatment, Vikström et al. (2016) tighten their bounds by employing the monotone treatment response assumption (Manski, 1997; and Manski and Pepper, 2000), a common shocks assumption that is employed in structural job search models (e.g., Meyer 1996), and by introducing a positive correlated outcomes assumption.

on the upper part of the employment spell distribution, the confidence intervals include zero for most percentiles above the median.

We complement our analysis of the JC effects on employment duration by estimating the effect of JC on the wages from those employment spells. This is an important feature of the proposed approach: it allows bounding not only the effects of training on employment duration but also on wages, since the bounds we employ address noncompliance and selection into employment (which are the identification problems for estimating wage effects). We find that, for the same subpopulation for which the effects on employment duration were estimated, JC increases their average wages in those employment spells by between 6.2 and 13.7 log points (6.4 and 14.7 percent), suggesting that JC not only helps these individuals to maintain their jobs longer, but also that those jobs are better paid. We also note that, in contrast to previous work that also analyzed wage effects of JC under noncompliance and sample selection (Chen and Flores, 2015), here we also consider quantile effects. Our results suggest heterogeneous effects along the wage distribution, with quantiles above the median ruling out zero effects more frequently than lower quantiles.

For unemployment, we consider the effect on the duration in weeks of the last complete unemployment spell before week 208 after randomization for those compliers who would experience a complete unemployment spell whether or not they enrolled in JC, a subpopulation representing about 39 percent of our non-Hispanics sample, and 56 percent of all compliers. In general, our results are not able to pin down the sign of this effect, as our estimated bounds (and confidence intervals) contain zero. For example, under our preferred set of assumptions, the average effect of JC on the unemployment duration outcome using our non-Hispanics sample is bounded between -10.5 log points (-10 percent) and 14.5 log points (15.6 percent).

Finally, our results indicate that the effects of interest are heterogeneous across the different demographic groups analyzed, where white males seem to benefit more from JC training relative to other demographic groups in terms of both longer employment spells and higher wages.

The remainder of the paper is organized as follows. Section 2 presents the econometric framework and assumptions used to construct bounds while addressing selection, noncompliance and censoring. Section 3 describes the JC program, the NJCS experimental data we use, and presents our empirical results. We present a discussion of important results in Section 4 and conclude in Section 5.

2 Econometric Framework and Assumptions

Consider a random sample of size N from a large population. Let $Z_i = z \in \{0, 1\}$ indicate whether unit i was randomly assigned to the treatment group ($Z_i = 1$) or to the control group ($Z_i = 0$). Let $T_i(z)$ denote the binary potential treatment (actually) received as a function of treatment assignment $Z_i = z$, that is, $T_i(1)$ and $T_i(0)$ represent the treatment participation status of unit i when randomly assigned to the treatment and control groups, respectively. Then, the observed treatment receipt indicator is $T_i = Z_i T_i(1) + (1 - Z_i) T_i(0) = t \in \{0, 1\}$. Given the nature of the outcomes of interest, we define the binary potential censoring $W_i(z, t)$ as a function of treatment assignment $Z_i = z$ and actual treatment receipt $T_i = t$. In our application, the observed censoring variable $W_i = w \in \{0, 1\}$ is an indicator of employment at the end of the observation period, and thus, the employment duration outcome will be censored if $W_i = 1$, while the unemployment duration outcome will be censored if $W_i = 0$. Lastly, we define the potential outcome $Y_i(z, t, w)$ as a function of the randomized treatment assignment, actual treatment receipt and censoring.

Following Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996, AIR hereafter), we use the potential treatment $T_i(z)$ to define the following four subpopulations based on their compliance behavior: the *compliers*, $c = \{i : T_i(0), T_i(1) = (0, 1)\}$; the *always-takers*, $a = \{i : T_i(0), T_i(1) = (1, 1)\}$; the *never-takers*, $n = \{i : T_i(0), T_i(1) = (0, 0)\}$; and the *defiers*, $d = \{i : T_i(0), T_i(1) = (1, 0)\}$. We start our analysis by focusing on compliance behavior and employing the following assumptions:

Assumption 1 Stable Unit Treatment Value Assumption (Rubin, 1978, 1980, 1990).

Assumption 2 Randomized Treatment Assignment, $Z_i \perp\!\!\!\perp \{T_i(z), W_i(z, t), Y_i(z, t, w)\}$ for $z, t, w \in \{0, 1\}$.

Assumption 1 implies that the potential outcomes for each unit i are unrelated to treatment status of other units. Assumption 2 is satisfied in our application. With imperfect compliance of the treatment assignment, not even under Assumptions 1 and 2 will the difference of average outcomes by treatment assignment yield an unbiased estimator of the average treatment effect (ATE). The following assumptions are necessary to define a causal effect of T_i on Y_i:

Assumption 3 Nonzero Average Causal Effect of Z_i on T_i, $E[T_i(1) - T_i(0)] \neq 0$.

Assumption 4 Individual-level Monotonicity of T_i in Z_i, $T_i(1) \geq T_i(0)$ *for all* i.

Assumption 5 Exclusion Restriction of Z_i, $Y_i(z, t, w) = Y_i(z', t, w) = Y_i(t, w)$ and $W_i(z, t) =$

Table 1: Pricipal Strata within Observed cells defined by Z_i, T_i and W_i.

		$Z_i = 0$		$Z_i = 1$	
		$T_i =$		$T_i =$	
		0	1	0	1
$W_i =$	0	*cNN, cNE, nNN*	*aNN*	*nNN*	*cNN, cEN, aNN*
	1	*cEE, cEN, nEE*	*aEE*	*nEE*	*cEE, cNE, aEE*

$W_i(z', t) = W_i(t)$ for all $z, z', t \in \{0, 1\}$ with $z \neq z'$.

Assumption 3 requires Z_i to have a non-zero effect on T_i. Assumption 4 states that there is no unit i that does the opposite of his/her random assignment, allowing us to rule out the *defier* subpopulation $d = \{i : T_i(0), T_i(1) = (1, 0)\}$. Finally, the first part of Assumption 5 states that any effect of the random assignment Z_i on the outcome Y_i must be via the effect of Z_i on the actual treatment receipt T_i. In the absence of any other econometric issue, Assumptions 1 to the first part of 5 were employed by Imbens and Angrist (1994) and AIR to show that Instrumental Variables estimators point identify the average treatment effect for the subpopulation of *compliers*, $E[Y(1) - Y(0)|c]$.

To address censoring, we use principal stratification (Frangakis and Rubin, 2002) to define subpopulations based on values for the potential censoring, now written as $W_i(z)$.[6] Given the censoring variable in our application (employment), we define the following four subpopulations: *always-employed*, $EE = \{i : W_i(0), W_i(1) = (1, 1)\}$; *employed only if assigned to the treatment group*, $NE = \{i : W_i(0), W_i(1) = (0, 1)\}$; *never-employed*, $NN = \{i : W_i(0), W_i(1) = (0, 0)\}$; and, *employed only if assigned to the control group*, $EN = \{i : W_i(0), W_i(1) = (1, 0)\}$. We proceed by combining the subpopulations based on the potential compliance behavior and censoring, that is $\{a, n, c, d\} \times \{EE, NE, NN, EN\}$. Out of 16 latent principal strata, based on Assumptions 1 to 4, we can eliminate *dEE, dNE, dNN and dEN* since these contain *defiers*; and based on the second part of Assumption 5, we can eliminate *aNE, aEN, nNE, and nEN* since for these strata there is some effect of the treatment assignment Z_i on censoring W_i that is not working through the actual treatment receipt T_i, in other words, given the latter part of Assumption 5, $W_i(1) = W_i(0)$ for noncompliers *a, n.*

In Table 1 we show the mixture of principal strata contained within observed cells defined by the values of Z_i, T_i and W_i, after employing Assumptions 1 to 5. It should be noted that depending on the outcome of interest, it is only possible to (partially) identify causal effects for one stratum, that is, without employing additional assumptions. For the employment spell outcome, a completed spell is observed when $W_i = 0$, which indicates that the individual is

[6]Defining subpopulations based on compliance behavior, as in Imbens and Angrist (1994) and AIR, is a special case of the principal stratification framework by Frangakis and Rubin (2002).

not employed at the end of the observation period, otherwise with $W_i = 1$ the actual spell is censored. Then, one can construct bounds for the effect of actual treatment T_i on the employment spell Y_i for *compliers* who are never employed at the end of the observation period, the cNN stratum. Conversely, for the unemployment outcome a full spell is only observed when $W_i = 1$, and one can bound the effect of T_i on Y_i for *compliers* who are always employed at the end of the observation period, the cEE stratum.[7] Then, in general, our focus is on bounding the effect of T_i on Y_i for the *compliers* whose full spells are always observed regardless of treatment assignment Z_i. First, we define the Complier Average Treatment Effect ($CATE$). The $CATE$ expressions for the outcome of employment and unemployment spells, superscript e and u respectively, are as follow

$$CATE^e \equiv E[Y_i(1)|cNN] - [Y_i(0)|cNN] \tag{1}$$

$$CATE^u \equiv E[Y_i(1)|cEE] - [Y_i(0)|cEE]. \tag{2}$$

Next, we discuss our identification strategy and the expressions for the respective bounds.

2.1 Bounds on Complier Treatment Effects Controlling for Censoring

We start by defining other necessary quantities. Let π_k denote the proportion of the principal stratum k in the population, that is, k could be any of the 8 different strata in Table 1. Let $p_{tw|z} \equiv Pr(T_i = t, W_i = w|Z_i = z)$, for $t, w, z \in \{0, 1\}$. In addition, let $P_{k|z}$ denote the distribution of the potential outcome for the k principal stratum, and define the distribution of each observed outcome Y_i for units with $(T_i, Z_i, W_i) = (t, z, w)$ as P_{tzw}. As shown on Table 1, with Assumptions 1 to 5 we point identify the strata proportions $\pi_{aNN} = p_{10|0}$, $\pi_{aEE} = p_{11|0}$, $\pi_{nNN} = p_{00|1}$, and $\pi_{nEE} = p_{01|0}$, which yield the following relationships,

$$\begin{aligned} \pi_{cNN} + \pi_{cNE} &= p_{00|0} - p_{00|1} \\ \pi_{cNN} + \pi_{cEN} &= p_{10|1} - p_{10|0} \\ \pi_{cEE} + \pi_{cEN} &= p_{01|0} - p_{01|1} \\ \pi_{cEE} + \pi_{cNE} &= p_{11|1} - p_{11|0} \end{aligned} \tag{3}$$

Then, under Assumptions 1 to 5, when $W_i = 0$ we analyze effects on employment spells since

[7] Note that for simplicity we use Y_i to denote both types of outcomes we consider, but we also make sure to distinguish between the analysis of employment and unemployment spells.

the outcome distributions $P_{t,z,0}$ are observed. It is also implied that the following relationships hold,

$$\frac{p_{00|0}P_{000} - p_{00|1}P_{010}}{p_{00|0} - p_{00|1}} = \frac{\pi_{cNN}}{p_{00|0} - p_{00|1}}P_{cNN|0} + (1 - \frac{\pi_{cNN}}{p_{00|0} - p_{00|1}})P_{cNE|0} \tag{4}$$

$$\frac{p_{10|1}P_{110} - p_{10|0}P_{100}}{p_{10|1} - p_{10|0}} = \frac{\pi_{cNN}}{p_{10|1} - p_{10|0}}P_{cNN|1} + (1 - \frac{\pi_{cNN}}{p_{10|1} - p_{10|0}})P_{cEN|1}. \tag{5}$$

In equation (4), the distribution to the left of the equality is obtained by subtracting the observed distribution P_{010} from P_{000} using $p_{00|1}/(p_{00|0} - p_{00|1})$ and $p_{00|0}/(p_{00|0} - p_{00|1})$ as respective weights. The latter operation implies subtracting the distribution of potential outcomes for nNN from $cNN + cNE + nNN$, where the result, in the right hand side of (4), is a mixture expressed as the weighted sum of cNN and cNE, with weights consistent with the first relationship in (3). An analogous procedure is used to obtain the relationship in equation (5).

Similarly, when analyzing unemployment, one observes full spells when $W_i = 1$. With the distributions $P_{t,z,1}$ now observed, the following relationships also hold based on Assumptions 1 to 5,

$$\frac{p_{01|0}P_{001} - p_{01|1}P_{011}}{p_{01|0} - p_{01|1}} = \frac{\pi_{cEE}}{p_{01|0} - p_{01|1}}P_{cEE|0} + (1 - \frac{\pi_{cEE}}{p_{01|0} - p_{01|1}})P_{cEN|0} \tag{6}$$

$$\frac{p_{11|1}P_{111} - p_{11|0}P_{101}}{p_{11|1} - p_{11|0}} = \frac{\pi_{cEE}}{p_{11|1} - p_{11|0}}P_{cEE|1} + (1 - \frac{\pi_{cEE}}{p_{11|1} - p_{11|0}})P_{cNE|1}. \tag{7}$$

Note that a testable implication of Assumptions 1 to 5 is that the left hand side of equations (4) to (7) must be valid probability distributions, that is, non-negative everywhere and integrating to one. Additional testable implications are that the quantities in the denominators in equations (4) to (7) should be non-negative, or, equivalently, the relationships in (3) imply that $p_{00|0} \geq p_{00|1}$, $p_{10|1} \geq p_{10|0}$, $p_{01|0} \geq p_{01|1}$, and $p_{11|1} \geq p_{11|0}$.

We now define the Complier Quantile Treatment Effects ($CQTE$) on the outcomes of employment and unemployment duration as,

$$CQTE^e(\alpha) \equiv q_{cNN|1}(\alpha) - q_{cNN|0}(\alpha) \tag{8}$$

$$CQTE^u(\alpha) \equiv q_{cEE|1}(\alpha) - q_{cEE|0}(\alpha), \tag{9}$$

where $q_{cNN|z}$ and $q_{cEE|z}$ are the α-quantiles of the distributions $P_{cNN|z}$ and $P_{cEE|z}$, respectively, with $\alpha \in (0,1)$. Along with the $CATE$ in (1) and (2), these are the parameters we partially identify in this paper.

Under Assumptions 1 to 5, Imai (2007) used the relationships in (3) to derive sharp bounds on π_{cNN}, which are the basis for bounding the $CATE^e$ in (1) and $CQTE^e(\alpha)$ in (8) when $W_i = 0$. Sharp bounds on π_{cNN} are given by $\max(0, p_{10|1} + p_{01|1} - p_{10|0} - p_{01|0}) \leq \pi_{cNN} \leq \min(p_{10|1} - p_{10|0}, p_{00|0} - p_{00|1})$. We show in the Internet Appendix that, under Assumptions 1 to 5, sharp bounds on π_{cEE} satisfy $0 \leq \pi_{cEE} \leq \min(p_{01|0} - p_{01|1}, p_{11|1} - p_{11|0})$, and these bounds are then the basis for bounding the $CATE^u$ in (2) and $CQTE^u(\alpha)$ in (9) when $W_i = 1$. The expressions for the bounds on $CATE$ and $CQTE$ involve minimum (min) and maximum (max) operators (see Imai, 2007), which cannot be addressed with standard estimation and inference procedures (Hirano and Porter, 2012). Another caveat of these bounds is that, in practice, estimates are too wide and often uninformative about the sign of the effects. Next, we consider two additional assumptions that will yield tighter bounds with a closed-form expression.

2.1.1 Bounds Under Monotonicity Assumption

Noting that $W_i(z) = W_i(t)$ for *compliers* we proceed by adding the following:

Assumption 6 Individual-level Monotonicity of W_i in T_i, $W_i(1) \geq W_i(0)$ for *compliers*.

Assumption 6 states that for every *complier* the effect of T_i on W_i is non-negative, that is, we rule out the stratum π_{cEN}. In our application, this assumption states that the effect of training on employment at the end of the observation period is non-negative for all the *compliers* of treatment assignment. Under Assumptions 1 to 6, after setting $\pi_{cEN} = 0$ in the second and third relationships in (3), the proportions of interest are point identify as

$$\pi_{cNN} = p_{10|1} - p_{10|0} \quad \text{and} \quad \pi_{cEE} = p_{01|0} - p_{01|1},$$

which in turn implies that the distributions in (5) and (6) belong to treated cNN and untreated cEE, respectively. Therefore, as shown below, we are able to identify one of the terms from the $CATE$ expressions in (1) and (2), and one of the terms from the $CQTE$ expressions in (8) and (9), whereas bounds need to be constructed to partially identify the unobserved counterfactuals. Following Imai (2007), we express the identified and partially identified quantities based on the following Q_{zw} distributions,

$$
(10) \qquad \begin{aligned} Q_{00} &\equiv \frac{p_{00|0}P_{000} - p_{00|1}P_{010}}{p_{00|0} - p_{00|1}}, & Q_{10} &\equiv \frac{p_{10|1}P_{110} - p_{10|0}P_{100}}{p_{10|1} - p_{10|0}}, \\ Q_{01} &\equiv \frac{p_{01|0}P_{001} - p_{01|1}P_{011}}{p_{01|0} - p_{01|1}}, \quad \text{and} & Q_{11} &\equiv \frac{p_{11|1}P_{111} - p_{11|0}P_{101}}{p_{11|1} - p_{11|0}}, \end{aligned}
$$

and their corresponding α-quantiles, $r_{zw}(\alpha) = \inf\{y : Q_{zw}[-\infty, y] \geq \alpha\}$, for $z, w \in \{0, 1\}$ and $\alpha \in (0, 1)$. For ease of exposition, below we separate the analysis by outcome type.

Effects on Employment Spells.

Under Assumptions 1 to 6, from the expressions for $CATE^e$ in (1) and $CQTE^e(\alpha)$ in (8), the term $E[Y_i(1)|cNN]$ and the α-quantile $q_{cNN|1}(\alpha)$ are identified, respectively, as

$$
(11) \qquad \overline{Y}_{10} = \int y \, dQ_{10}
$$

$$
(12) \qquad r_{10}(\alpha),
$$

which leaves the respective counterfactuals $E[Y_i(0)|cNN]$ and $q_{cNN|0}(\alpha)$ as not point identified. Note that from the empirical distribution in (4), with π_{cNN} point identified under Assumptions 1 to 6, we can construct a lower (upper) bound on $E[Y_i(0)|cNN]$ and $q_{cNN|0}(\alpha)$ by placing all individuals that belong to the cNN stratum at the bottom (top) $\frac{p_{10|1}-p_{10|0}}{p_{00|0}-p_{00|1}}$ portion in the probability distribution expressed by Q_{00}, which implies placing all individuals in cNE at the top (bottom) of the distribution. As a result, the lower and upper bounds for $E[Y_i(0)|cNN]$ are given by

$$
(13) \qquad L_{cNN,0} = \int y \, dL_{cNN|0}
$$

$$
(14) \qquad U_{cNN,0} = \int y \, dU_{cNN|0},
$$

with distributions $L_{cNN|0}$ and $U_{cNN|0}$ defined, respectively, as

$$
L_{cNN|0}[-\infty, y] \equiv \begin{cases} \frac{Q_{00}[-\infty,y]}{\pi_{cNN|0}} & \text{if} \quad y < r_{00}(\pi_{cNN|0}) \\ 1 & \text{if} \quad y \geq r_{00}(\pi_{cNN|0}) \end{cases}
$$

$$
U_{cNN|0}[-\infty, y] \equiv \begin{cases} 0 & \text{if} \quad y < r_{00}(1 - \pi_{cNN|0}) \\ \frac{Q_{00}[-\infty,y]-1+\pi_{cNN|0}}{\pi_{cNN|0}} & \text{if} \quad y \geq r_{00}(1 - \pi_{cNN|0}), \end{cases}
$$

where $\pi_{cNN|0} = \frac{p_{10|1}-p_{10|0}}{p_{00|0}-p_{00|1}}$ is the proportion of individuals that belong to the cNN stratum within the cell $\{Z_i = 0, T_i = 0, W_i = 0\}$. An analogous procedure is used to construct the following lower and upper bounds on $q_{cNN|0}(\alpha)$:

$$q^l_{cNN,0}(\alpha) = r_{00}(\alpha\pi_{cNN|0}) \tag{15}$$

$$q^u_{cNN,0}(\alpha) = r_{00}(1-(1-\alpha)\pi_{cNN|0}). \tag{16}$$

Then, under Assumptions 1 to 6, we use the identified $E[Y_i(1)|cNN]$ in (11), and the bounds in (13) and (14) for the counterfactual $E[Y_i(0)|cNN]$, to partially identify the Complier Average Treatment Effect on employment spells in (1), such that $LB_{CATE^e} \leq CATE^e \leq UB_{CATE^e}$, where

$$LB_{CATE^e} = \overline{Y}_{10} - U_{cNN,0} \tag{17}$$

$$UB_{CATE^e} = \overline{Y}_{10} - L_{cNN,0} \tag{18}$$

Similarly, we use the identified $q_{cNN|1}(\alpha)$ in (12), and the bounds in (15) and (16) for the counterfactual $q_{cNN|0}(\alpha)$, to partially identify the Complier Quantile Treatment Effect on employment spells in (8), yielding $LB_{CQTE^e(\alpha)} \leq CQTE^e(\alpha) \leq UB_{CQTE^e(\alpha)}$, where

$$LB_{CQTE^e(\alpha)} = r_{10}(\alpha) - q^u_{cNN,0}(\alpha) \tag{19}$$

$$UB_{CQTE^e(\alpha)} = r_{10}(\alpha) - q^l_{cNN,0}(\alpha). \tag{20}$$

Following Imai (2007 and 2008), one can show that bounds in (17, 18) and (19, 20) are sharp.

Effects on Unemployment Spells.

An analogous procedure is used to bound the effects on unemployment duration, where a full spell is now observed when $W_i = 1$. Therefore, under Assumptions 1 to 6, from the expressions for $CATE^u$ in (2) and $CQTE^u(\alpha)$ in (9), we respectively identify the terms $E[Y_i(0)|cEE]$ and $q_{cEE|0}(\alpha)$ as,

$$\overline{Y}_{01} = \int y\, dQ_{01} \tag{21}$$

$$r_{01}(\alpha). \tag{22}$$

Lower (upper) bounds on the counterfactuals $E[Y_i(1)|cEE]$ and $q_{cEE|1}(\alpha)$ are calculated by placing all individuals that belong to the cEE stratum at the bottom (top) $\frac{p_{01|0}-p_{01|1}}{p_{11|1}-p_{11|0}}$ portion in the distribution Q_{11}, which is shared with individuals from the cNE stratum. Then,

$$L_{cEE,1} = \int y \, dL_{cEE|1} \tag{23}$$

$$U_{cEE,1} = \int y \, dU_{cEE|1} \tag{24}$$

are the corresponding lower and upper bounds for $E[Y_i(1)|cEE]$, where the distributions $L_{cEE|1}$ and $U_{cEE|1}$ are defined as,

$$L_{cEE|1}[-\infty, y] \equiv \begin{cases} \frac{Q_{11}[-\infty,y]}{\pi_{cEE|1}} & \text{if} \quad y < r_{11}(\pi_{cEE|1}) \\ 1 & \text{if} \quad y \geq r_{11}(\pi_{cEE|1}) \end{cases}$$

$$U_{cEE|1}[-\infty, y] \equiv \begin{cases} 0 & \text{if} \quad y < r_{11}(1-\pi_{cEE|1}) \\ \frac{Q_{11}[-\infty,y]-1+\pi_{cEE|1}}{\pi_{cEE|1}} & \text{if} \quad y \geq r_{11}(1-\pi_{cEE|1}), \end{cases}$$

and $\pi_{cEE|1} = \frac{p_{01|0}-p_{01|1}}{p_{11|1}-p_{11|0}}$ is the proportion of individuals that belong to the cEE stratum within the cell $\{Z_i = 1, T_i = 1, W_i = 1\}$. Analogously, the following are the lower and upper bounds on $q_{cEE|1}(\alpha)$:

$$q^l_{cEE,1}(\alpha) = r_{11}(\alpha\pi_{cEE|1}) \tag{25}$$

$$q^u_{cEE,1}(\alpha) = r_{11}(1-(1-\alpha)\pi_{cEE|1}). \tag{26}$$

As with the previous outcome, under Assumptions 1 to 6, we use the identified term in (21), and the bounds in (23) and (24) to partially identify the $CATE^u$ in (2), as $LB_{CATE^u} \leq CATE^u \leq UB_{CATE^u}$, where

$$LB_{CATE^u} = L_{cEE,1} - \overline{Y}_{01} \tag{27}$$

$$UB_{CATE^u} = U_{cEE,1} - \overline{Y}_{01} \tag{28}$$

Then, to partially identify $CQTE^u(\alpha)$ in (9) we use the identified term in (22), and the bounds in (25) and (26), such that $LB_{CQTE^u(\alpha)} \leq CQTE^u(\alpha) \leq UB_{CQTE^u(\alpha)}$, where

$$LB_{CQTE^u(\alpha)} = q^l_{cEE,1}(\alpha) - r_{01}(\alpha) \tag{29}$$

$$UB_{CQTE^u(\alpha)} = q^u_{cEE,1}(\alpha) - r_{01}(\alpha). \tag{30}$$

It can be shown that the bounds in (27, 28) and (29, 30) are sharp bounds, the proof is omitted since it is a straightforward extension of the bounds by Imai (2007), who considered a case in which the outcome was observed if $W_i = 0$.

2.1.2 Bounds Under Stochastic Dominance Assumption

Tighter bounds for the causal effects of interest can be constructed, under Assumptions 1 to 6, by adding the following assumption:

Assumption 7 Stochastic Dominance Across Compliers Strata. **7.a**. For the outcome of employment spells: $P_{cNN|z}[-\infty, y] \geq P_{cNE|z}[-\infty, y]$. **7.b**. For the outcome of unemployment spells: $P_{cEE|z}[-\infty, y] \geq P_{cNE|z}[-\infty, y]$.

This additional assumption states that, regardless of treatment, the potential outcome of one stratum at any rank of the outcome distribution is at least as large as that of other stratum. In particular, 7.a. implies that, at the end of the observation period, the last full employment spell for compliers never employed at week 208, cNN, are equal or smaller than the spells for compliers whose employment probability is affected positively by training, cNE, and this holds at any point of the respective distributions of the outcome. On the other hand, 7.b. implies that, at the end of the observation period, the last full unemployment spell for compliers always employed at week 208, cEE, are equal or smaller than the spells for compliers whose employment probability is affected positively by training, cNE.

Tightening Bounds for the Effects on Employment Spells.

To understand the tightening power of adding stochastic dominance to the previous set of assumptions, we focus on the relationship in (4). With Assumption 7.a. the empirical distribution Q_{00}, which is equivalent to the right hand side in (4), is now a lower bound for the distribution of the cNN stratum, $P_{cNN|0}$, since it contains a mixture of distributions for the cNN and, the stochastic dominant, cNE strata. As a result, the upper bounds for the counterfactuals $E[Y_i(0)|cNN]$ and $q_{cNN|0}(\alpha)$ are now tighter than (14) and (16), respectively, and calculated as:

$$\overline{Y}_{00} = \int y \, dQ_{00} \tag{31}$$

$$r_{00}(\alpha). \tag{32}$$

Then, under Assumptions 1 to 7.a, tighter bounds for the parameters of interest are: $LB^*_{CATE^e} \leq CATE^e \leq UB_{CATE^e}$ and $LB^*_{CQTE^e(\alpha)} \leq CQTE^e(\alpha) \leq UB_{CQTE^e(\alpha)}$, where the respective upper bounds remain as in (18) and (20), and the tighter lower bounds are given by,

$$LB^*_{CATE^e} = \overline{Y}_{10} - \overline{Y}_{00} \tag{33}$$

$$LB^*_{CQTE^e(\alpha)} = r_{10}(\alpha) - r_{00}(\alpha). \tag{34}$$

Under Assumptions 1 to 7.a, the bounds based on (33) and (34) can be shown to be sharp; note that they are obtained by using the tighter upper bounds in (31) and (32) to respectively replace the quantities $U_{cNN,0}$ in (14) and $q^u_{cNN,0}(\alpha)$ in (16), which were used in constructing the sharp lower bounds in (17) and (19) under Assumptions 1 to 6. A formal proof is a straightforward extension of a proof presented in Imai (2007), who derived sharp bounds in a case were the distribution of the stratum of interest, here cNN, stochastically dominates the other stratum, cNE in the present context.

Tightening Bounds for the Effects on Unemployment Spells.

We now focus on the distribution Q_{11}, which is equivalent to the right hand side in (7). Adding Assumption 7.b makes Q_{11} a lower bound for the distribution of potential outcomes for the cEE stratum, $P_{cEE|1}$, since it contains a mixture of distributions for two strata, the cEE and the stochastic dominant cNE. As a result, relative to (24) and (26), the new upper bounds for the counterfactuals $E[Y_i(1)|cEE]$ and $q_{cEE|1}(\alpha)$ are tighter and calculated as:

$$\overline{Y}_{11} = \int y \, dQ_{11} \tag{35}$$

$$r_{11}(\alpha). \tag{36}$$

It follows that under Assumptions 1 to 7.b, $LB_{CATE^u} \leq CATE^u \leq UB^*_{CATE^u}$ and $LB_{CQTE^u(\alpha)} \leq CQTE^u(\alpha) \leq UB^*_{CQTE^u(\alpha)}$, where the respective lower bounds remain as in (27) and (29), and the tighter upper bounds are given by,

$$UB^{*}_{CATE^{u}} = \overline{Y}_{11} - \overline{Y}_{01} \tag{37}$$

$$UB^{*}_{CQTE^{u}(\alpha)} = r_{11}(\alpha) - r_{01}(\alpha). \tag{38}$$

With the addition of Assumption 7.b, bounds based on (37) and (38) can be shown to be sharp since they were obtained by using the tighter upper bounds in (35) and (36) to respectively replace the quantities $U_{cEE,1}$ in (24) and $q^{u}_{cEE,1}(\alpha)$ in (26), which were used to construct the sharp upper bounds in (28) and (30) under Assumptions 1 to 6. A formal proof is a straightforward extension of a proof presented in Imai (2007).

2.2 Estimation

We use the indicator function $1[Y_i \leq \tilde{y}]$ to identify the cumulative distribution function (cdf) of the observed outcome Y_i evaluated at $\tilde{y}$, that is, $\hat{P}_{tzw}(\tilde{y})$ for $t, z, w \in \{0, 1\}$. Then, an entire cdf, $\hat{P}_{tzw}$, is constructed by using M different values of $\tilde{y}$ spanning the support of the observed outcome. We also employ the following sample analogs for the proportions $p_{tw|z}$,

$$\begin{aligned}
\hat{p}_{00|0} &= \frac{\sum_{i=1}^{n}(1-T_i)\cdot(1-W_i)\cdot(1-Z_i)}{\sum_{i=1}^{n}(1-Z_i)}, & \hat{p}_{00|1} &= \frac{\sum_{i=1}^{n}(1-T_i)\cdot(1-W_i)\cdot Z_i}{\sum_{i=1}^{n}Z_i}, \\
\hat{p}_{10|1} &= \frac{\sum_{i=1}^{n}T_i\cdot(1-W_i)\cdot Z_i}{\sum_{i=1}^{n}Z_i}, & \hat{p}_{10|0} &= \frac{\sum_{i=1}^{n}T_i\cdot(1-W_i)\cdot(1-Z_i)}{\sum_{i=1}^{n}(1-Z_i)}, \\
\hat{p}_{01|0} &= \frac{\sum_{i=1}^{n}(1-T_i)\cdot W_i\cdot(1-Z_i)}{\sum_{i=1}^{n}(1-Z_i)}, & \hat{p}_{01|1} &= \frac{\sum_{i=1}^{n}(1-T_i)\cdot W_i\cdot Z_i}{\sum_{i=1}^{n}Z_i}, \\
\hat{p}_{11|1} &= \frac{\sum_{i=1}^{n}T_i\cdot W_i\cdot Z_i}{\sum_{i=1}^{n}Z_i}, & \hat{p}_{11|0} &= \frac{\sum_{i=1}^{n}T_i\cdot W_i\cdot(1-Z_i)}{\sum_{i=1}^{n}(1-Z_i)}.
\end{aligned}$$

Subsequently, the empirical cdf's $\hat{P}_{tzw}$ and the weights given by $\hat{p}_{tw|z}$ are used to construct estimates for the distributions Q_{zw} in (10), such that

$$\begin{aligned}
\hat{Q}_{00} &= \frac{\hat{p}_{00|0}\hat{P}_{000} - \hat{p}_{00|1}\hat{P}_{010}}{\hat{p}_{00|0} - \hat{p}_{00|1}}, & \hat{Q}_{10} &= \frac{\hat{p}_{10|1}\hat{P}_{110} - \hat{p}_{10|0}\hat{P}_{100}}{\hat{p}_{10|1} - \hat{p}_{10|0}}, \\
\hat{Q}_{01} &= \frac{\hat{p}_{01|0}\hat{P}_{001} - \hat{p}_{01|1}\hat{P}_{011}}{\hat{p}_{01|0} - \hat{p}_{01|1}}, \quad \text{and} & \hat{Q}_{11} &= \frac{\hat{p}_{11|1}\hat{P}_{111} - \hat{p}_{11|0}\hat{P}_{101}}{\hat{p}_{11|1} - \hat{p}_{11|0}}.
\end{aligned}$$

Finally, the estimates $\hat{Q}_{zw}$ are used to compute the expected values and inverted to get the α-quantiles needed to estimate the bounds for $CATE$ and $CQTE$, respectively.

3 Analyzing the Effects of Job Corps Training on Employment and Unemployment Duration

Here we employ the bounds described in the previous section to assess the effect of Job Corps (JC) training on the duration of employment and unemployment spells, but first we briefly discuss the JC program and data, and also provide a preliminary—albeit naive—analysis of the effect of JC on employment and unemployment durations.

3.1 Job Corps, Data and Preliminary Analysis

The JC program was established in 1964 under the Economic Opportunity Act, and today is a key pillar of the Workforce Innovation and Opportunity Act (WIOA), signed in 2014. The program is administer by the US Department of Labor (DOL) and is America's largest and most comprehensive no-cost education and job training program. In line with the design of WIOA, the goal of the JC program is to help economically disadvantaged young people, ages 16 to 24, improve the quality of their lives by enhancing their labor market opportunities and educational skill set, which is achieved through the offering of academic instruction, career technical training, residential living, health care, counseling, and job placement assistance. In a typical year, about 60,000 eligible students enroll in one of the 125 JC centers located nationwide, where participants typically reside.[8] Due to its comprehensive nature, the annual cost of the program ascends to over \$1.6 billion (DOL, Office of Inspector General report in 2013).

Being the nation's largest job training program, the evaluation of the JC effectiveness is of public interest. During the mid nineties, the DOL 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 contiguous states and the District of Columbia, and were randomly assigned to treatment and control groups. From a randomly selected research sample of 15,386 first time eligible applicants, 9,409 were assigned to the treatment group and the remainder 5,977 to the control group, during the sample intake period from November 1994 to February 1996. After recording their data through a baseline interview for both treatment and control groups, a series of follow up interviews were conducted at weeks 52, 130, and 208 after randomization (Schochet

[8] 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. For more information see Schochet et al. (2001).

et al., 2001).

Our sample is restricted to individuals who have non-missing values for weekly earnings and hours worked for every week after random assignment.[9] These restrictions were employed by Lee (2009) and Blanco et al. (2013a). In addition, the sample is restricted to individuals with information on actual enrollment, captured by a binary indicator of whether the individual was ever enrolled in JC during the 208 weeks after randomization. Chen and Flores (2015) employed the same set of restrictions to analyze the effects of enrolling in JC on wages. In order to increase the likelihood of Assumption 6, individual level monotonicity, we focus on the sub-sample that excludes Hispanics.[10] Finally, we employ the NJCS design weights throughout the analysis, since different subgroups in the population had different probabilities of being included in the research sample (for details on NJCS design weights see Schochet, 2001).

As shown at the bottom of Table 2, the Full sample has 9,094 individuals: 5,496 and 3,598 in the randomized treatment and control groups, respectively. The sub-sample sample of interest, Non-Hispanics, has 7,531 individuals: 4,554 and 2,977 assigned to treatment and control groups, respectively.[11] In the first row of Table 2 we report the extent of noncompliance in our data based on the Enrollment indicator. In both samples, roughly about 74 percent of the individuals in the treatment group actually enrolled in JC within the 208 weeks after randomization. During the same period, the proportion of control group individuals that enrolled in JC were 4.4 and 4.7 percent for the Full and Non-Hispanic samples, respectively.

Given the noncompliance in these samples, 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.[12] The estimates reported on the third and fourth rows under the columns labeled Difference correspond to the ITT effects for the outcome in logs. Being offered participation in JC increases the lengths of employment at week 208 after randomization by 6 to 7 log points, and unemployment by about 8 to 10 log points; however, only unemployment effects are statistically significant at a 5 percent level in both samples. An

[9] We implicitly assume—as do studies cited in this paragraph—that the missing values are "missing completely at random".

[10] More details on why we exclude Hispanic can be learned from Blanco et al. (2013a). We can estimate appropriate bounds for the effects of interest in Hispanics without employing Assumptions 6, however, as noted in Section 2.1, these bounds would likely be wide an uninformative.

[11] Using a 5 percent significance level, all but 3 out of 26 selected pre-treatment covariate averages do not differ statistically between treatment and control groups, in the Full and Non-Hispanic samples (see Internet Appendix). The latter is an expected result given randomization. We also note that after excluding Hispanics from the Full sample the magnitudes of average pre-treatment covariates do not change in a meaningful way.

[12] For studies analyzing ITT effects of the JC program see, for example, Schochet et al. (2001), Lee (2009), Zhang et al. (2009), Flores-Lagunes et al. (2010), and Blanco et al. (2013a, 2013b), where the papers by Lee (2009) and Blanco et al. (2013a, 2013b) bound ITT effects of JC on wages.

alternative well-known estimator we consider is the Local Average Treatment Effect ($LATE$), which addresses noncompliance and we report in the following two rows. For both outcomes, the average treatment effects for the compliers, $LATE$, is about 3 to 4 percentage points higher in magnitude than the estimated ITT. As with the ITT, estimates of $LATE$ are only significant at a 5 percent level for the outcome of log of unemployment duration (in weeks) at week 208 after randomization. Of course, these ITT and $LATE$ estimates ignore censoring and sample selection, and thus are biased. Note that treatment assignment (Z) significantly increases the probability of being employed at week 208 after randomization (our censoring indicator in the second row) by about 4.0 and 4.9 percentage points for the Full and Non-Hispanic samples, respectively.

A preliminary (naive) analysis of quantile treatment effects is presented in Table 3. We employ two estimators that are analogous to the average treatment effect estimators in Table 2. First, we consider the comparison at different points of the outcome distributions by treatment assignment (Koenker and Bassett, 1978). These Quantile Treatment Effect (QTE) estimates for the 25^{th}, 50^{th} and 75^{th} percentiles are presented in the upper half of the table. In both samples, employment length at week 208 after randomization is affected positively by treatment assignment, but the only statistically significant effect, at a 10 percent level, is for the 25^{th} percentile with a magnitude of 15.4 log points. It is also important to note that effects become smaller in magnitude at higher percentiles of the outcome distribution. Except for the Full sample's 25^{th} percentile, where the estimated effect is zero, unemployment length is affected positively by treatment assignment at the considered quantiles, but the only significant effect, at a 5 percent level, is the 18 log points increase at the 25^{th} percentile for the non-Hispanic sample. These effects also diminish in magnitude at higher percentiles. Second, to control for noncompliance we employ the Instrumental Variable Quantile Treatment Effect ($IVQTE$) estimator by Abadie, Angrist and Imbens (2002). Estimates based on $IVQTE$ are reported in the bottom half of the table. The estimated quantile effects of JC on the length of employment at week 208 after randomization for *compliers* remains positive, with notable increases in magnitude observed in the Non-Hispanic sample, relative to the QTE estimates. As before, the only statistically significant effect is estimated for the 25^{th} percentile, this time at a 5 percent level. Except for the zero effect at the 25^{th} percentile, in both samples we note that the lengths of unemployment at week 208 after randomization for *compliers* are affected positively by enrollment in JC, where estimates are significant at a 10 and 5 percent level for the 25^{th} and 50^{th} percentiles, respectively, in the Non-Hispanic sample. In general, effects based on $IVQTE$ are larger in magnitude than those based on QTE. Both of these estimators, however, ignore censoring and are biased. We now present our main estimates and analysis, which are based on the nonparametric bounds estimators, discussed in section 2, that control for noncompliance, selection, and censoring.

3.2 Main Results

3.2.1 Bounds for $CATE$

Table 4 presents our estimated bounds for the $CATE$ on the uncensored outcomes of employment and unemployment durations in weeks of the last complete spell before week 208 after randomization, in logs, along with their respective Imbens and Manski (IM, 2004) confidence intervals. The table also shows the estimated stratum proportions and quantities used in estimating the bounds under Assumptions 1 to 6, and under Assumptions 1 to 7. For brevity, we focus on the Non-Hispanic sample estimates, which are presented in column 1.[13] In addition, we analyze the extent of heterogeneous impacts on other demographic groups of interest in columns 2 to 9. In the Non-Hispanic sample, the largest estimated stratum proportion corresponds to the cEE stratum, which accounts for about 39 percent of the population (and 56 percent of all compliers), followed by the cNN stratum, with an estimated proportion of about 26 percent (and 37 percent of the compliers). Hence, our analysis of employment and unemployment spells is relevant to the two largest strata, which together account for about 65 percent of the population and 93 percent of all the compliers. While the latter statement holds, in general, for the other demographic groups we consider, there are some interesting differences in the actual estimated proportion magnitudes. For example, the cEE stratum represents about 50 percent in the White Males sample, while the cNN stratum represents 30 percent in the Black Males sample. Importantly, in every sub-sample, the remainder of estimated stratum proportions are all highly statistically significant (standard errors not reported here).

Estimated Bounds for $CATE$ on Employment Spells. Focusing on the Non-Hispanic sample, under Assumptions 1 to 6 the estimated lower and upper bounds for the $CATE$ on the logarithm of the duration in weeks of the last full employment spell completed before week 208 after randomization are -0.384 and 0.465, respectively. While negative effects are not ruled out, these estimated bounds cover a larger positive region and it is plausible that effects are larger than those reported based on the ITT and $LATE$ point estimates, which where smaller than 0.11. A similar observation is made based on the estimated bounds for the sub-samples

[13] We note that main results for the Full sample are remarkably similar to those for the sample of Non-Hispanics, as it was the case in the preliminary analysis in Section 3.1. For brevity, results for the Full sample are relegated to the internet appendix.

Whites, Males, Females and Black Females, while the estimated bounds for the samples of White Females and Black Males cover a larger negative region. Interestingly, the estimated lower bound for White Male *compliers* is the only one that rules out large negative effects, with an estimated 2.3 log points reduction in the employment spells due to JC training. Nevertheless, bounds are uninformative about the sign of the effect for all sub-samples analyzed.

One drawback of Assumption 6 is that it implies a perfect negative correlation between being employed and employment duration in constructing the lower bound, which is highly unlikely. For example, this is evident in the Non-Hispanic sample, as it positions at the top of the mixed distribution in $\hat{Q}_{00}$ the 83.9 percent representing the cNN stratum, as shown by the estimated $\pi_{cNN|0}$. The benefit of adding Assumption 7 is that it rules out this implausible negative relation to construct a tighter lower bound. The estimated lower bound for $CATE$ under Assumptions 1 to 7 for the Non-Hispanic sample is now positive, and statistically significant based on the 90 percent IM confidence interval. That is, the duration (in weeks) of the last completed employment spell before week 208 after randomization for Non-Hispanic *compliers* who would experience a full employment spell regardless of treatment assignment increased, on average, by at least 10.7 log points due to JC training, with this effect being statistically significant with 10 percent confidence. For all other demographic groups, the lower bounds are positive, although the effects are statistically insignificant based on the IM confidence intervals. The exception is the corresponding stratum for White Males, for whom JC training has a positive and statistically significant average effect on our employment duration outcome, with the increase in duration being at least 33 log points and at most 55 log points.

Analysis of Wages During Employment Spells.

To complement the analysis on average employment spells, we bound the effects of JC training on the average wage during the observed employment duration.[14] These results are reported in Table 5, and for brevity we only discuss estimates based on Assumptions 1 to 7, at the bottom of the top panel. For Non-Hispanics, Whites, Blacks, Males and White Males the effects of JC training on the average wage during the employment spells and for the stratum considered above (cNN) is positive and statistically significant, based on the IM confidence interval for the lower bound. For these groups, the lower bound for the effect ranges from 3.8 (Blacks) to 15.6 log points (White Males), while the upper bound is consistent with increases in wages that vary from 10 (Blacks) to 19.5 log points (White Males). For Females, White Females, Black Males

[14]To bound the effects of JC training on the average wage during the observed employment duration, we employed the same bounds used to analyze employment spells but with the outcome of average wages during the spell. These bounds are appropriate since they essentially deal with noncompliance and the selection that allows one to observe these average wages, under the same set of assumptions.

and Black Females the estimated lower bounds are positive but statistically insignificant.

These results suggest that training in JC increases the average duration in employment for the Non-Hispanic sample and their average wages while employed, and that these effects are relatively high. Thus, on average, these individuals are not only able to keep their jobs longer, but these jobs are better paid. Except for White Males, we cannot rule out zero average effects on the duration of employment for the other demographic groups analyzed, but average wages while employed for Whites, Blacks and Males are significantly higher for *cNN compliers* training in JC. The complementary analysis also reinforces the finding that White Males are benefiting relatively more.

Estimated Bounds for $CATE$ on Unemployment Spells.

Back to Table 4, where the lower panel shows that, under Assumptions 1 to 6, the estimated lower and upper bounds for the $CATE$ on the logarithm of the duration in weeks of the last full unemployment spell completed before week 208 after randomization for the Non-Hispanic sample are -0.105 and 0.433, respectively. Like the estimates for employment spells under Assumptions 1 to 6, the estimated bounds for $CATE$ on unemployment spells are not indicative of a plausible direction for the effect of interest. Moreover, one cannot rule out large increases in unemployment duration for *cEE compliers* that trained in JC. For the other demographic groups analyzed the qualitative result is the same, namely, the estimated bounds are not informative about the sign of the effect of JC training on unemployment spells for *cEE compliers*, and positive effects can potentially be large.

In contrast to the analysis on employment duration effects, the drawback of Assumption 6, when analyzing effects on unemployment duration, is that it implies a perfect positive correlation between being employed and unemployment spells in constructing the upper bound, and this is highly unlikely. Note that for the Non-Hispanic sample, under Assumptions 1 to 6, the upper bound is constructed by positioning at the top of the mixed distribution in $\hat{Q}_{11}$ the 88.8 percent representing the cEE stratum, as shown by the estimated $\pi_{cEE|1}$. When adding Assumption 7 we rule out this implausible positive relation and get a tighter upper bound. For Non-Hispanics, the estimated upper bound for $CATE$ for our unemployment duration outcome under Assumptions 1 to 7 is now 0.145. While this tighter upper bound does not rule out positive effects, at least it is able to rule out large increases due to JC training in the unemployment spells for *cEE compliers*. In contrast to the naive estimates in Table 2, the estimated lower bound for $CATE$ is suggestive of a plausible large reduction in unemployment spells, in the order of 10.5 log points, while the estimated upper bound is quite close to the point identified

$LATE$ (i.e., an increase of about 14 log points). Relative to the Non-Hispanic sample, results for the other demographic groups analyzed are very similar, that is, uninformative about the sign of the effect.

3.2.2 Bounds for $CQTE$

In going beyond the analysis of average effects, we present our estimated bounds for the $CQTE$ on the duration (in weeks) of the last completed employment and unemployment spell before week 208 after randomization in logs, along with their respective IM confidence intervals, in Figure 1 for the Non-Hispanic sample and Figures 2 to 5 for the other demographic groups of interest. Actual numerical results for this portion of the analysis are relegated to the internet appendix.

Estimated Bounds for $CQTE$ on Employment Spells. Similarly to the estimated bounds for average effects, in the Non-Hispanic sample, estimates of the lower and upper bounds for the $CQTE$ on employment spells under Assumptions 1 to 6, shown in Figure 1(a), are not informative about the sign of the effect in any of the analyzed percentiles. However, a few points are worth mentioning. An interesting trend based on the upper bound alone is that effects increase as one moves up the distribution of employment spells, this is more evident in percentiles beyond the median, while for percentiles below the median the upper bound magnitudes are very similar to the estimated upper bound for $CATE$. Based on lower bounds, a similar trend is observed since reductions in employment spells become smaller in magnitude as one moves up the distribution. Relative to the bounds on $CATE$, the bounds at the median are narrower, estimates are -0.26 and 0.35.

Results after we add Assumption 7 are presented in Figure 1(b). The estimated lower bounds become tighter and positive in all the percentiles we consider. However, based on the 90 percent IM confidence intervals, only the lower bounds at the 25^{th} and 40^{th} percentiles are statistically significant. Similarly to the estimated bounds for $CATE$, the lower bound at the median is positive, however, not statistically significant. Relative to the $IVQTE$ estimates in Table 2, the upper bounds in all percentiles do not rule out much larger effects, however, the lower bounds are very similar in magnitudes and slightly smaller at the 25^{th} and 50^{th} percentile, and almost 3 percentage points larger at the 75^{th} percentile. The previous observation about having significant effects in lower quantiles holds true based on our bounds analysis.

For brevity, we only discuss estimated bounds for $CQTE$ on Employment Spells after employing Assumptions 1 to 7 in our analysis of other demographic sub-samples, which are summarized in Figure 3. We leave Figure 2—which shows the corresponding bounds under Assumptions 1 to 6—to illustrate the identifying power of Assumption 7, and note the sizable reductions in the lower bounds, which go from negative to positive in most cases. Focusing on Figure 3, we note that in a couple of lower percentiles in the Whites sample we identify a significant positive effect. No lower bound is statistically significant in the Blacks sample. The effect at the median in the Males sample is statistically significant, where JC training increases employment duration for *cNN compliers* by at least 19 log points and up to 41 log points. There are a few percentiles under the median where the lower bound is indicative of a positive JC training effect on the employment duration of Females. Relative to other demographic groups, large and significant effects are observed in the White Males sample in percentiles surrounding the median. In contrast, no lower bound is statistically significant in the White Females sample and the effect could plausibly be negative in percentiles above the median. Interestingly, we observe that results in the Black Males sample are qualitatively similar to those of White Females, while effects are positive and significant at the 25^{th} and other higher percentiles when analyzing the observed employment duration distribution of Black Females. By and large, bounds suggest that the positive effects of JC are more pronounced in the White Males sample.

Analysis of Wages During Employment Spells.

A distributional analysis of average wages during the last completed employment spell before week 208 after randomization for *cNN compliers* is presented in the upper graphs from Figure 6 for the Non-Hispanic sample. Focusing on the estimated bounds after employing Assumptions 1 to 7, Figure 6 (b), we note that significant effects are found in most percentiles above the first quartile, that is, the lower bounds are positive and significant based on the IM confidence intervals. Furthermore, effects are relatively larger beyond the 75^{th} percentile. In combination with the estimates for employment duration, these results are very insightful since they suggest that JC training has a positive impact in the bottom part of the distribution of employment duration and also a positive impact in almost all of the distribution beyond the 25^{th} percentile of average wages during the employment spells.

Figure 7 presents the analysis on average wages during the employment spells for the other demographic groups considered, under Assumptions 1 to 7. A very similar pattern as the one noted for the Non-Hispanic sample can be seen when analyzing the samples of Whites and White Males, where significant positive impacts on employment duration were estimated in some of the lower percentiles while significant positive impacts on the average wage during these spells are larger in higher percentiles of the distribution. We also find that wage impacts are larger in higher percentiles for the samples of Males and, to a lesser extent, Blacks and Black Males. White Females is the only sample where we did not rule out insignificant wage effects throughout the wage distribution, while some of the percentiles analyzed in the samples of Females and Black Females are indicative of significant positive wage effects.

Estimated Bounds for $CQTE$ on Unemployment Spells.

We turn back our attention to the sample of Non-Hispanics. Estimated bounds for $CQTE$ on our unemployment duration outcome of interest for the *cEE compliers* under Assumptions 1 to 6 are presented in Figure 1 (c). All the estimated upper bounds at the different percentiles we analyze are positive, so one cannot rule out increases in unemployment spells. One interesting trend, based on the upper bound, is that magnitudes of the effect decrease as one moves upward in the distribution. Many of the lower bounds at percentiles below the median are 0, ruling out reductions in unemployment duration due to JC training in this portion of the distribution. In percentiles above the median, lower bounds are also consistent with relatively large reductions in unemployment duration. Compared to the bounds for $CATE$, the estimated bounds for the effect at the median are slightly narrower, but with a lower bound, -0.17, that is consistent with larger potential reductions in unemployment duration and an upper bound, 0.36, consistent with a smaller potential increase.

In Figure 1 (d) we add Assumption 7. As expected, the estimated upper bounds in all the percentiles considered become much tighter but still positive, which does not allow one to rule out a positive effect on unemployment spells. It now becomes clearer that the lower bound is negative and much larger in magnitude than the positive upper bounds for percentiles above the median. Importantly, these results cast doubts on the point estimates based on $IVQTE$ reported in Table 2, where it was suggested that effects where positive and statistically significant at the 25^{th} and 50^{th} percentiles.

Our analysis on the other demographic sub-samples is presented in Figure 4, where bounds are constructed employing Assumptions 1 to 6, and Figure 5 where we add Assumption 7. Similarly to the analysis of employment spells, going from Figure 4 to 5 illustrates the identifying power of Assumption 7, although it is generally observed that reductions in the upper bound are not sizable enough to rule out increases in unemployment duration. For brevity, we only discuss interesting estimated bounds for $CQTE$ on Unemployment Spells after employing Assumptions 1 to 7, reported in Figure 5. The results observed in the Non-Hispanic sample distribution in Figure 1 (d) are similar to those in the Whites, Blacks and Males samples in Figure 5 (a), (b) and (c). In percentiles above the median for Females, lower and upper bounds are very similar in magnitudes, that is, neither large negative nor positive effects are ruled out. White Males are largely unaffected in percentiles below the median, where most bounds point identified 0, and relatively large effects are ruled out above the median. For White Females, the upper bound at the median and at the 60^{th} percentile are consistent with a reduction in the unemployment spell, however these bounds are not statistically significant based on the IM confidence intervals. For Black Males, upper bounds in the 75^{th} and 80^{th} percentiles are also negative but insignificant. Finally, we note that Black Females' results are similar to those reported for the Non-Hispanic sample.

4 Discussion of Results

Evidence presented above suggests that JC training has a positive average effect on the duration of the last complete employment spell before week 208 after randomization for cNN *compliers*, a stratum representing about 26 percent of the Non-Hispanic sample (and 37 percent of all compliers).[15] Under the preferred set of assumptions, the estimated bounds suggest that the statistically significant increase in employment spells could have been between 10.7 to 46.5 log points for the Non-Hispanic sample. In addition, we provide evidence suggesting that the effect in question is heterogeneous along the outcome distribution. This is the case because the effect at some lower percentiles we study are positive and statistically distinguishable from zero, as suggested by the estimated lower bounds and the respective IM confidence interval. Larger effects can potentially happen on the upper part of the employment spell distribution, however, a zero effect cannot be ruled out for upper percentiles. In a supplemental analysis, we find that JC training not only increases the duration in employment but also the average wages while employed. We also find heterogeneous results in our analysis of demographic sub-groups, where it is clear that White Males are benefited the most by training in JC. It is important to point out that other studies in the literature analyzing JC have also found that the program has a relatively larger positive impact for whites (see, for example, Flores-Lagunes et al., 2010, who suggested that the program may be shielding whites from adverse local economic conditions).

Our estimated bounds for the average effect of JC training on cEE compliers' observed unemployment spells is relevant to about 39 percent of the individuals in the Non-Hispanic sample, that would be the most important stratum in terms of size. Based on the full set of assumptions, we find that JC training could potentially increase unemployment spells by up

[15] As previously discussed, the strata of interest are the most important ones in terms of size, and this is also true in the other demographic groups we analyze. We should note that the conclusions based on the Non-Hispanic sample largely apply to the Full sample too.

to 14.5 percent, on the other hand, the estimated lower bound suggests a potential reduction of up to 10.5 percent, on average. Once again, our analysis of quantiles sheds light on the heterogeneity of training effects. We do not rule out an increase in unemployment spells, however, estimates for the higher percentiles we study are more consistent with a reduction, while the effects at lower percentiles rule out reductions. In the demographic sub-groups studied there is a great deal of heterogeneity, but bounds are not informative about the sign of the effect. Importantly, the bounds analysis casts doubts on conclusion based on point estimators, where it was suggested that JC increases unemployment duration significantly. The general trend suggests that those with a largest potential benefit are located in the upper part of the distribution of unemployment spells. These individuals are the most disadvantaged from a stratum that has better labor market outcomes (at baseline), relative to individuals whose employment, and hence the observability of unemployment spells, was affected positively by treatment assignment (cNE).

5 Conclusions

With a yearly cost of about $1.6 billion, Job Corps is the nation's largest job training program targeting disadvantaged youth. As such, providing evidence about the program's effectiveness is of extreme importance for policy making purposes. Here, we use experimental data to analyze the impact of the JC program on the distribution of the duration of employment and unemployment spells of participants. To our knowledge, we are the first study to shed light on the program's impact on these important duration outcomes.

To achieve our goal, we employ principal stratification to define our parameters of interest and propose an approach to the estimation problem at hand that allows us to use nonparametric bounds for average and quantile treatment effects analogous to those derived in Imai (2007) to address two identification problems (sample selection and noncompliance), to also address an additional complication present in our setting, censoring. As a result, the proposed approach allows the bounds we employ to address three identification issues: sample selection, noncompliance, and censoring; three pervasive problems that are present even in randomized evaluations and that can lead to bias and misleading conclusions if they are ignored. In contrast to conventional parametric techniques, the nonparametric bounds used herein employ relatively weak assumptions to control for these identification problems.

We present evidence suggesting that training in Job Corps increases average employment duration in the last completed spell before week 208 after randomization for cNN compliers by at least 10.7 log points and up to 46.5 log points. Our results are unable to pin down the sign of the average effect on the unemployment duration in the last completed spell before week 208 after randomization for cEE compliers. The quantile results reflect heterogeneous effects and strengthen our conclusions based on the average effects. We complement the analysis of spells and report positive average wage effects during those employment spells. We also report that effects differ by demographic groups, where White Males seem to be benefiting relatively more in terms of larger effects on both employment duration and wages.

6 References

Abadie, A., Angrist, J., and Imbens, G. 2002. “Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings” Econometrica, 70: 91-117.

Abbring J., and Van den Berg, G. 2003. “The Nonparametric Identification of Treatment Effects in Duration Models.” Econometrica, 71(5): 1491-1517.

Angrist, J., Imbens, G., and Rubin, D. 1996. “Identification of Causal Effects Using Instrumental Variables.” Journal of the Statistical Association, 91: 444-455.

Blanco, G., Flores, C., and Flores-Lagunes, A. 2013a. “Bounds on Average and Quantile Treatment Effects of Job Corps Training on Wages.” Journal of Human Resources, 48(3): 659-701.

Blanco, G., Flores, C., and Flores-Lagunes, A. 2013b. “The Effects of Job Corps Training on Wages of Adolescents and Young Adults.” American Economic Review P&P, 103(3): 418-422.

Blundell, R., and Powell, J. 2007. “Censored Regression Quantiles with Endogenous Regressors.” Journal of Econometrics, 141: 6583.

Chen, X. and Flores, C. 2015.“Bounds on Treatment Effects in the Presence of Sample Selection and Noncompliance: The Wage Effects of Job Corps.” Journal of Business and Economic Statistics, 33 (4): 523-540.

Chen, X., Flores, C., and Flores-Lagunes, A. (2017), ”Going beyond LATE: Bounding Average Treatment Effects of Job Corps Training,” Mimeo, Department of Economics, California Polytechnic State University at San Luis Obispo.

Chernozhukov, V., Fernandez-Val, I., and Kowalski, A. 2015. “Quantile Regression with

Censoring and Endogeneity." Journal of Econometrics, 186: 201-221.

Chernozhukov, V., and Hong, H. 2002. "Three-Step Censored Quantile Regression and Extramarital Affairs." Journal of the American Statistical Association, 97: 872-882.

Dehejia, R. 2005. "Practical Propensity Score Matching: A Reply to Smith and Todd." Journal of Econometrics, 125: 355-364.

Dehejia, R., and Wahba, S. 1999. "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." Journal of the American Statistical Association, 94: 1053-1062.

Dehejia, R., and Wahba, S. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." The Review of Economics and Statistics, 84(1): 151-161.

Eberwein, C., Ham, J., and LaLonde, R. 1997. "The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women: Evidence from Experimental Data." Review of Economic Studies, 64: 655-682.

Eberwein, C., Ham, J., and LaLonde, R. 2002. "Alternative Methods of Estimating Program Effects in Event History Models." Labour Economics, 9: 249-278.

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

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

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.

Frandsen, B. 2015. "Treatment Effects with Censoring and Endogeneity." Journal of the American Statistical Association, 110: 17451752.

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

Frölich, M., and Melly B. 2013. "Unconditional Quantile Treatment Effects under Endogeneity." Journal of Business and Economic Statistics, 31: 346-357.

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

Ham, J., and LaLonde, R. 1996. “The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training.” Econometrica 64: 175-205.

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 Singer. 1984. “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data.” Econometrica, 52(2): 271-320.

Hirano, K., and Porter, J. 2012. “Impossibility Results for Nondifferentiable Functionals.” Econometrica, 80 (4): 1769-1790.

Imai, K. 2007. “Identification Analysis for Randomized Experiments with Noncompliance and Truncation-by-Death.” Technical Report, Department of Politics, Princeton University.

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.

Koenker, R., and Bassett, G. 1978. “Regression Quantiles.” Econometrica 46: 33-50.

LaLonde, R. 1986. “Evaluating the Econometric Evaluations of Training Programs with Experimental Data.” The American Economic Review, 76(4): 604-620.

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

Manski, C. 1997. “Monotone Treatment Response.” Econometrica, 65: 1311-1334.

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

Meyer, B. 1996. “What have We Learned from the Illinois Reemployment Bonus Experiment?” Journal of Labor Economics, 14: 26-51.

Powell, J. 1986. “Censored Regression Quantiles.” Journal of Econometrics, 32: 143-155.

Rubin, D. 1978. “Bayesian Inference for Causal Effects.” The Annals of Statistics, 6: 34-58.

Rubin, D. 1980. Comment on “Randomization Analysis of Experimental Data: The Fisher Randomization Test,” by D. Basu, Journal of American Statistical Association, 75: 591-593.

Rubin, D. 1990. Comment: “Neyman (1923) and Causal Inference in Experiments and Observational Studies.” Statistical Science, 5: 472-480.

Sant’Anna, P. 2016. “Program Evaluation with Right Censored Data.” Mimeo.

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.

Smith, J., and Todd, P. 2005. “Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators?” Journal of Econometrics, 125: 305-353.

US Department of Labor. 2013. http://www.dol.gov/dol/topic/training/jobcorps.html.

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

Vikström, J., Ridder, G., and Weidner, M. 2016. “Bounds on Treatment Effects on Transitions.” Cemmap working paper CWP17/16.

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 2: Summary Statistics of Post-Treatment Variables and Preliminary Analysis for the Full and Non-Hispanic Samples, by Treatment Status

	Full Sample				Non-Hispanic Sample			
	Treatment $Z_i = 1$	Control $Z_i = 0$	Difference	Std.Err.	Treatment $Z_i = 1$	Control $Z_i = 0$	Difference	Std.Err.
At week 208:								
Enrollment	0.738	0.044	0.695	0.007 ***	0.737	0.047	0.689	0.008 ***
Employed	0.612	0.572	0.040	0.007 ***	0.613	0.564	0.049	0.011 ***
Intention to Treat								
Effect (ITT) at week 208:			ITT				ITT	
Log employment spell			0.060	0.047			0.072	0.051
Log unemployment spell			0.081	0.038 **			0.098	0.042 **
Local Average								
Treatment Effects								
(LATE) at week 208:			LATE				LATE	
Log employment spell			0.089	0.069			0.108	0.076
Log unemployment spell			0.114	0.053 **			0.139	0.059 **
Observations	5496	3598			4554	2977		

Note: *, **, and *** denote statistical significance difference in means, Diff., at a 90, 95 and 99 percent confidence level. Standard errors for the difference in means are reported under Std.Err. Computations use design weights.

Table 3: Preliminary Quantile Treatment Analysis for the Full and Non-Hispanic Samples

	Full Sample α-quantiles			Non-Hispanic Sample α-quantiles		
	0.25	0.50	0.75	0.25	0.50	0.75
Quantile Treatment Effect (QTE)						
Log employment spell	0.154 *	0.061	0.050	0.154*	0.061	0.050
	(0.088)	(0.057)	(0.062)	(0.091)	(0.059)	(0.066)
Log unemployment spell	0.000	0.057	0.039	0.182 **	0.057	0.059
	(0.076)	(0.063)	(0.048)	(0.078)	(0.072)	(0.061)
Intrumental Variable QTE (IVQTE)						
Log employment spell	0.154	0.061	0.050	0.288 **	0.125	0.074
	(0.124)	(0.077)	(0.074)	(0.131)	(0.078)	(0.082)
Log unemployment spell	0.000	0.118	0.101	0.223*	0.182**	0.125
	(0.138)	(0.084)	(3.850)	(0.14)	(0.098)	(0.082)

Note: Robust standard errors for QTE and bootstrapped standard errors for IVQTE are in parentheses. *, **, and *** denote statistical significance at a 90, 95 and 99 percent confidence level. Computations use design weights.

Table 4: Bounds for the Complier Average Treatment Effect on the Log of Length of Employment and Unemployment for Non-Hispanic Samples

	Non-Hispanic	Whites	Blacks	Males	Females	White Males	White Females	Black Males	Black Females
π_{cEE}	0.386	0.445	0.357	0.426	0.338	0.508	0.345	0.377	0.336
π_{aEE}	0.018	0.022	0.015	0.024	0.010	0.032	0.006	0.016	0.014
π_{nEE}	0.160	0.194	0.144	0.138	0.188	0.172	0.230	0.115	0.176
π_{cNN}	0.255	0.186	0.288	0.247	0.264	0.171	0.209	0.299	0.276
π_{cNE}	0.049	0.040	0.053	0.041	0.057	0.022	0.066	0.052	0.055
π_{aNN}	0.030	0.028	0.030	0.032	0.027	0.031	0.022	0.030	0.029
π_{nNN}	0.103	0.086	0.113	0.092	0.116	0.064	0.122	0.111	0.114
Analysis of Employment Spells									
$\pi_{cNN\|0}$	0.839	0.823	0.845	0.856	0.823	0.888	0.760	0.853	0.835
$\bar{Y}_{10}$	2.712	2.931	2.644	2.725	2.694	3.030	2.797	2.616	2.676
$\bar{Y}_{00}$	2.605	2.741	2.565	2.622	2.588	2.699	2.786	2.609	2.517
$L_{cNN,0}$	2.247	2.367	2.224	2.301	2.202	2.478	2.248	2.279	2.167
$U_{cNN,0}$	3.096	3.224	3.055	3.069	3.116	3.053	3.399	3.080	3.038
Bounds under Assumptions 1 to 6									
$[LB_{CATE^e}, UB_{CATE^e}]$	[-0.384, 0.465]	[-0.293, 0.563]	[-0.411, 0.421]	[-0.344, 0.424]	[-0.422, 0.492]	[-0.023, 0.552]	[-0.602, 0.549]	[-0.463, 0.337]	[-0.362, 0.509]
(90% IM confidence intervals)	(-0.539, 0.607)	(-0.595, 0.874)	(-0.602, 0.591)	(-0.556, 0.610)	(-0.639, 0.707)	(-0.391, 0.919)	(-1.054, 1.009)	(-0.717, 0.567)	(-0.651, 0.779)
Bounds adding Assumption 7									
$[LB^*_{CATE^e}, UB_{CATE^e}]$	[0.107, 0.465]	[0.190, 0.563]	[0.079, 0.421]	[0.104, 0.424]	[0.107, 0.492]	[0.331, 0.552]	[0.011, 0.549]	[0.007, 0.337]	[0.159, 0.509]
(90% IM confidence intervals)	(0.009, 0.607)	(-0.013, 0.874)	(-0.039, 0.591)	(-0.021, 0.610)	(-0.039, 0.707)	(0.080, 0.919)	(-0.307, 1.009)	(-0.155, 0.567)	(-0.033, 0.779)
Analysis of Unemployment Spells									
$\pi_{cEE\|1}$	0.888	0.918	0.871	0.911	0.856	0.959	0.839	0.880	0.860
$\bar{Y}_{01}$	2.655	2.501	2.759	2.617	2.709	2.440	2.635	2.774	2.740
$\bar{Y}_{11}$	2.800	2.633	2.879	2.739	2.891	2.557	2.795	2.844	2.919
$L_{cEE,1}$	2.550	2.444	2.597	2.535	2.578	2.449	2.445	2.578	2.599
$U_{cEE,1}$	3.088	2.837	3.216	2.966	3.268	2.661	3.184	3.150	3.292
Bounds under Assumptions 1 to 6									
$[LB_{CATE^u}, UB_{CATE^u}]$	[-0.105, 0.433]	[-0.057, 0.336]	[-0.162, 0.457]	[-0.082, 0.349]	[-0.131, 0.559]	[-0.009, 0.221]	[-0.189, 0.549]	[-0.196, 0.376]	[-0.141, 0.552]
(90% IM confidence intervals)	(-0.214, 0.548)	(-0.240, 0.508)	(-0.307, 0.619)	(-0.215, 0.483)	(-0.314, 0.765)	(-0.170, 0.389)	(-0.562, 0.939)	(-0.386, 0.569)	(-0.367, 0.801)
Bounds adding Assumption 7									
$[LB_{CATE^u}, UB^*_{CATE^u}]$	[-0.105, 0.145]	[-0.057, 0.133]	[-0.162, 0.120]	[-0.082, 0.122]	[-0.131, 0.182]	[-0.009, 0.117]	[-0.189, 0.160]	[-0.196, 0.070]	[-0.141, 0.179]
(90% IM confidence intervals)	(-0.214, 0.224)	(-0.240, 0.270)	(-0.307, 0.227)	(-0.215, 0.216)	(-0.314, 0.319)	(-0.170, 0.260)	(-0.562, 0.458)	(-0.386, 0.209)	(-0.367, 0.349)
Observations	7531	2339	4545	4251	3280	1504	835	2358	2187

Note: Bootstrapped standard errors to compute the Imbens and Manski (IM, 2004) confidence intervals are based on 1000 replications. Computations use design weights.

Table 5: Bounds for the Complier Average Treatment Effect on Ln(Wages) During Employment spells for Non-Hispanic Samples

	Non-Hispanic	Whites	Blacks	Males	Females	White Males	White Females	Black Males	Black Females
Analysis of Wages									
During Employment Spells:									
Bounds under Assumptions 1 to 6									
$[LB_{CATE^e}, UB_{CATE^e}]$	[-0.070, 0.137]	[-0.027, 0.191]	[-0.096, 0.100]	[-0.049, 0.130]	[-0.112, 0.121]	[0.023, 0.195]	[-0.112, 0.134]	[-0.079, 0.093]	[-0.114, 0.114]
(90% IM confidence intervals)	(-0.101, 0.173)	(-0.096, 0.271)	(-0.132, 0.142)	(-0.091, 0.177)	(-0.158, 0.177)	(-0.075, 0.299)	(-0.217, 0.254)	(-0.129, 0.147)	(-0.169, 0.183)
Bounds adding Assumption 7									
$[LB^{*}_{CATE^e}, UB_{CATE^e}]$	[0.062, 0.137]	[0.120, 0.191]	[0.038, 0.100]	[0.072, 0.130]	[0.044, 0.121]	[0.158, 0.195]	[0.044, 0.134]	[0.037, 0.093]	[0.038, 0.114]
(90% IM confidence intervals)	(0.032, 0.173)	(0.530, 0.271)	(0.003, 0.142)	(0.034, 0.177)	(-0.002, 0.177)	(0.071, 0.299)	(-0.063, 0.254)	(-0.009, 0.147)	(-0.015, 0.183)
Observations	7531	2339	4545	4251	3280	1504	835	2358	2187

Note: Bootstrapped standard errors to compute the Imbens and Manski (IM, 2004) confidence intervals are based on 1000 replications. Computations use design weights.

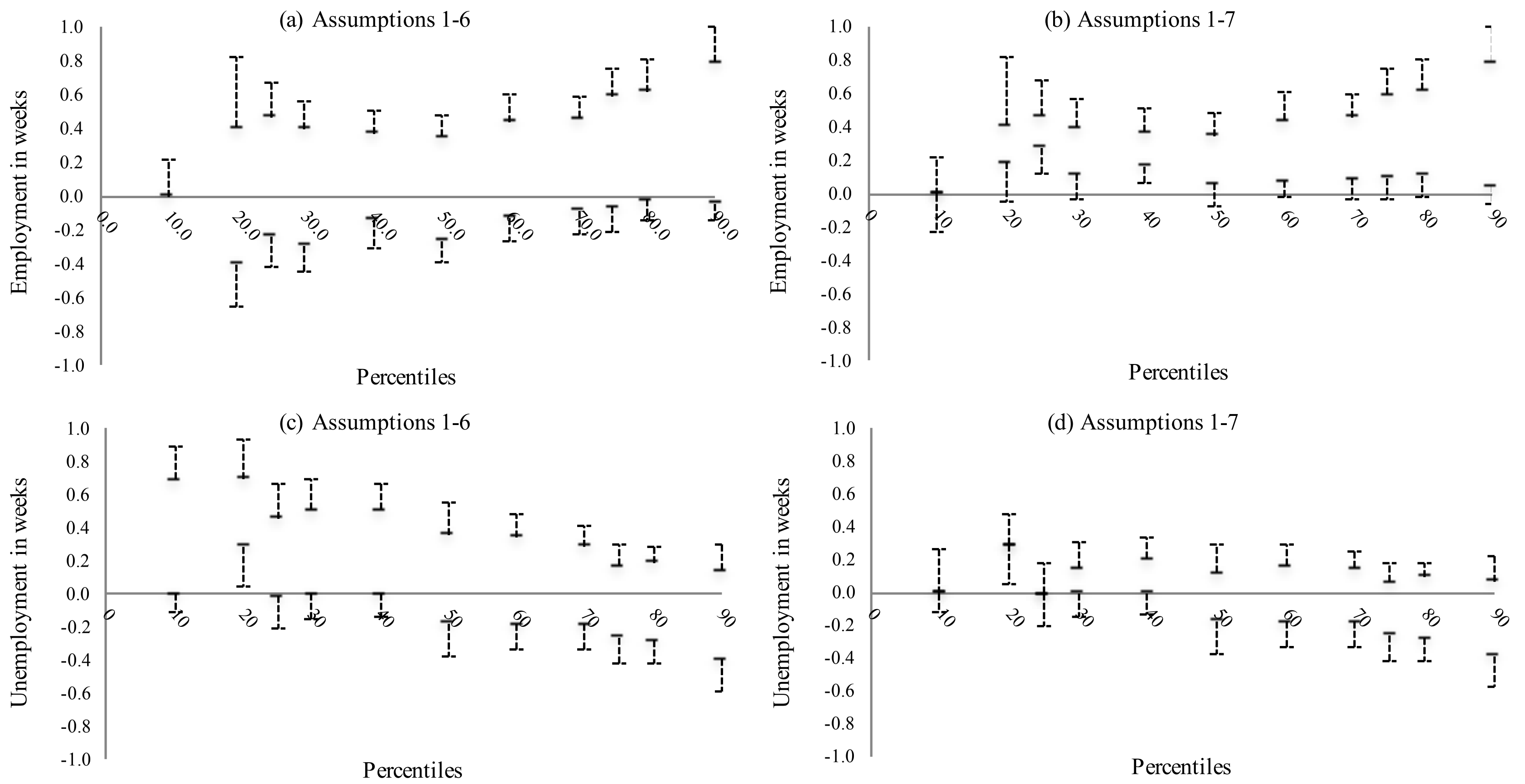

Figure 1. Bounds for CQTE on log of Employment (top graphs) and Unemployment (bottom graphs) spells for the Non-Hispanic sample. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

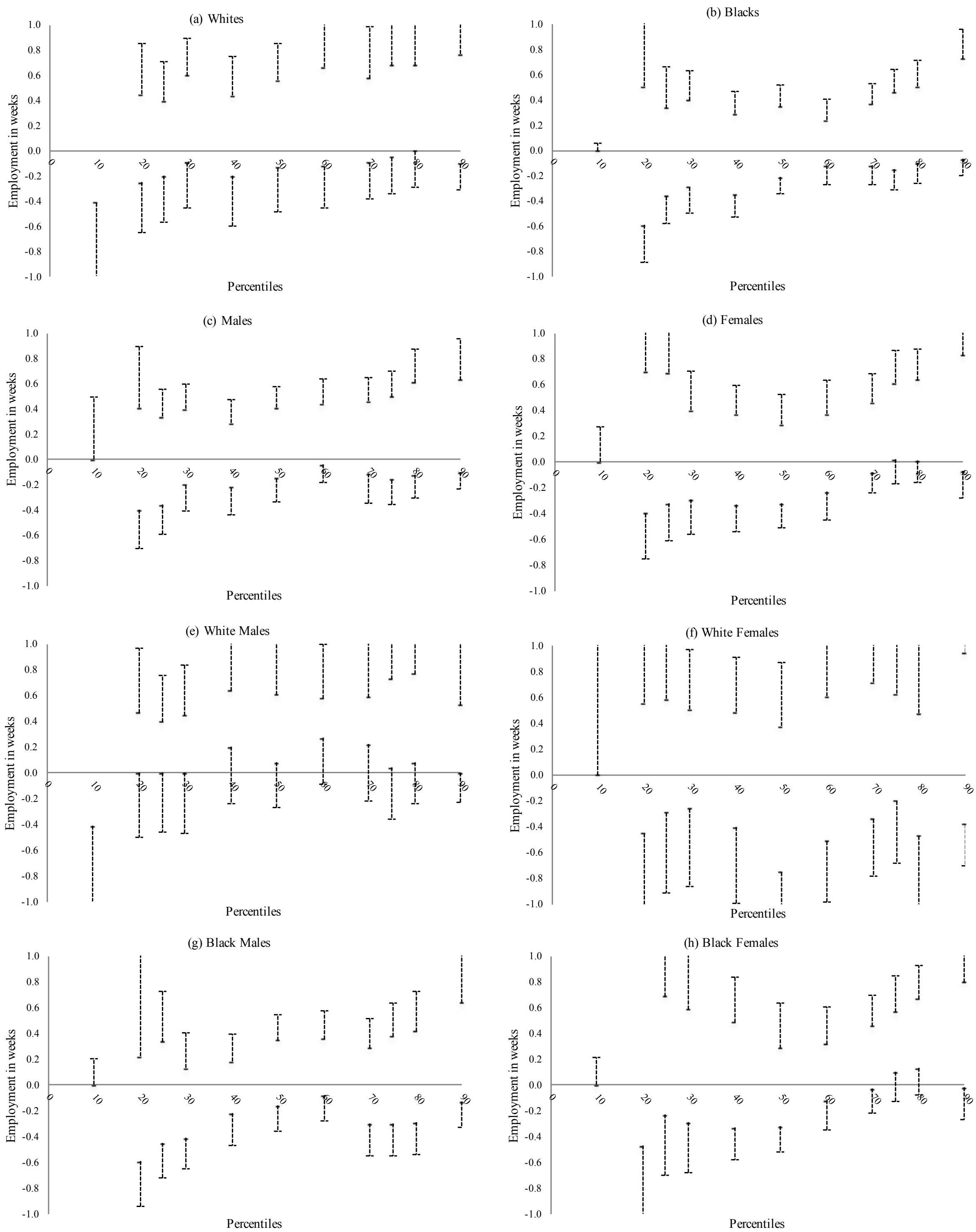

Figure 2. Bounds for CQTE on the log of Employment spells under Assumptions 1 to 6 for different demographic sub-samples. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

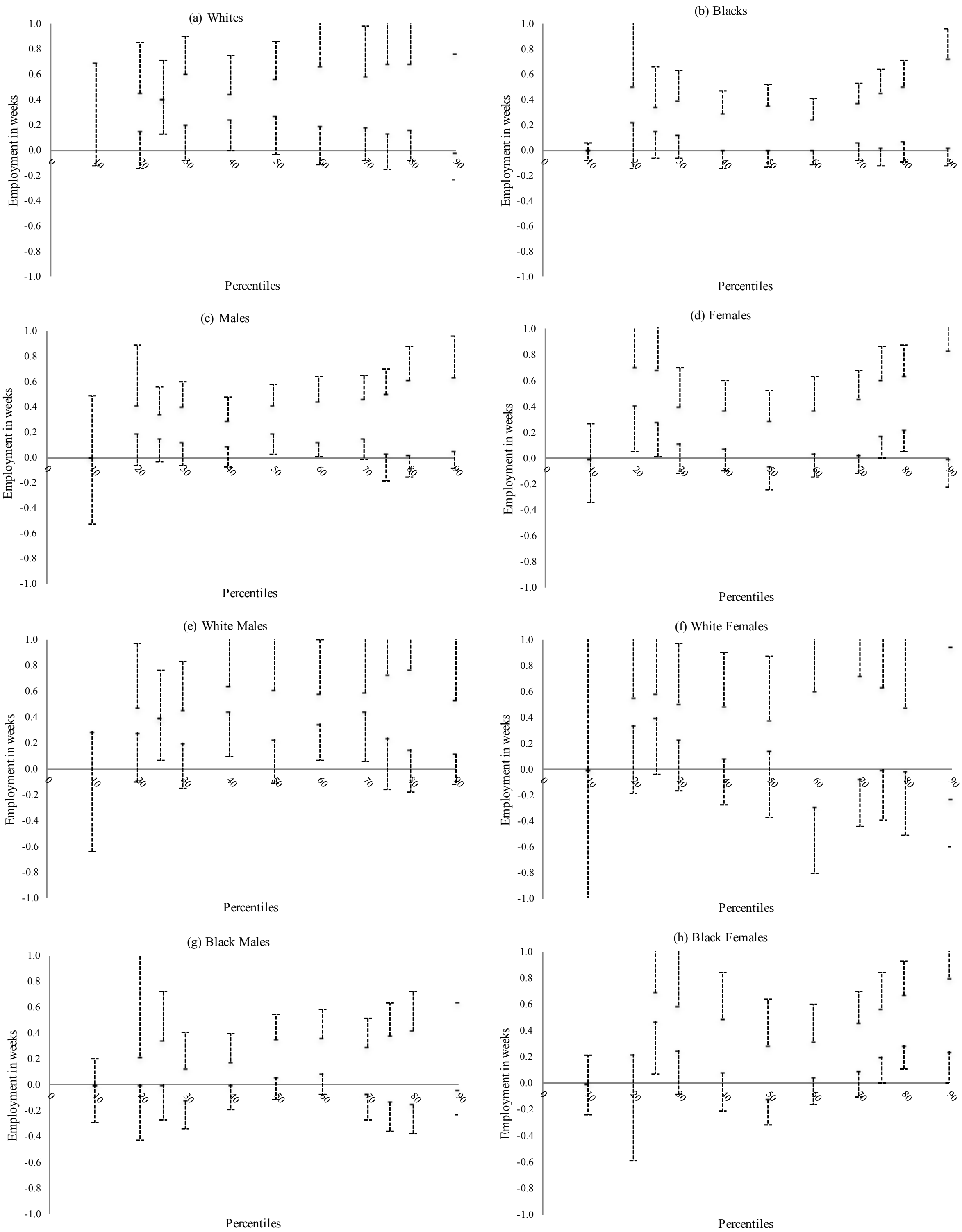

Figure 3. Bounds for CQTE on the log of Employment spells under Assumptions 1 to 7 for different demographic sub-samples. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

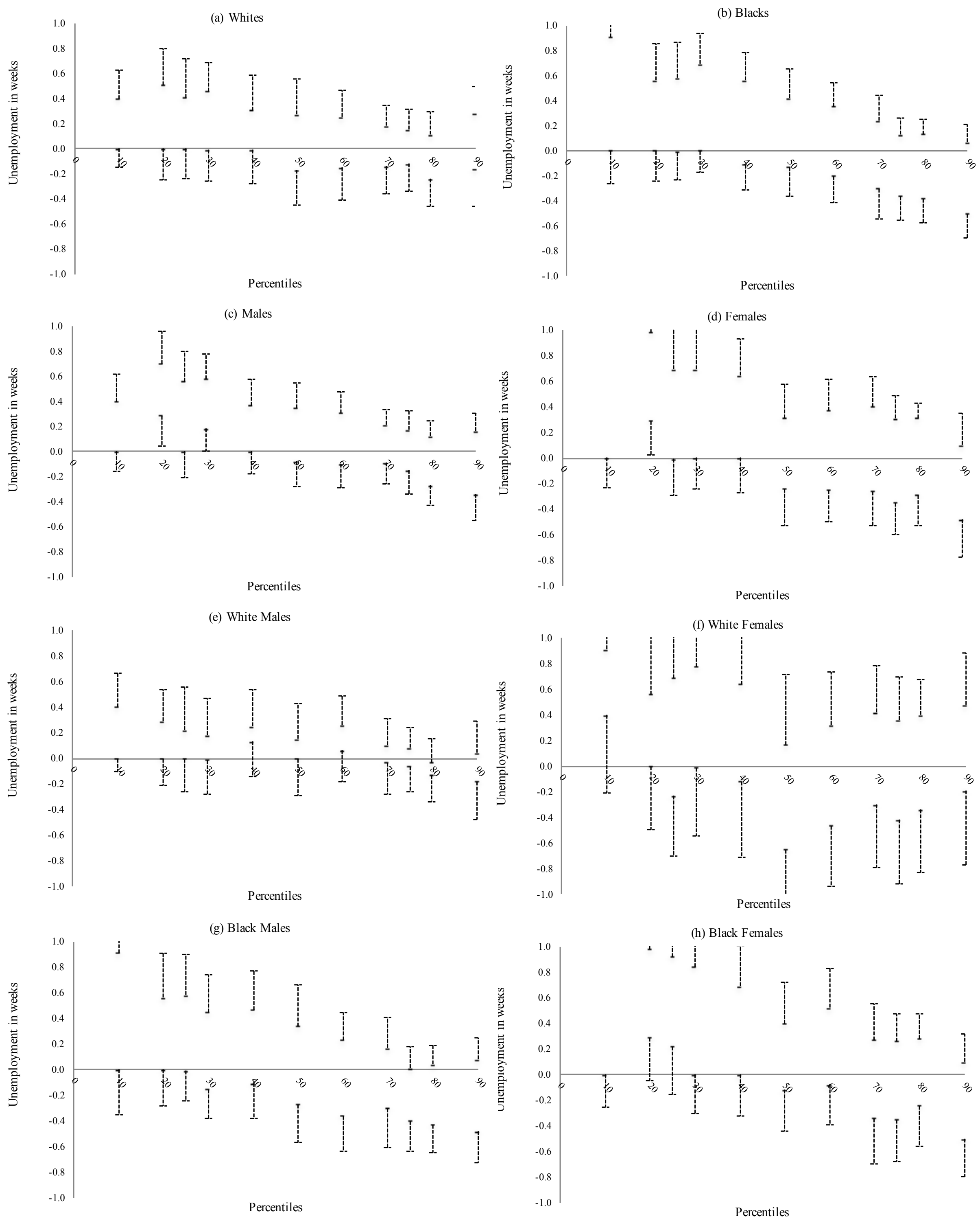

Figure 4. Bounds for CQTE on the log of Unemployment spells under Assumptions 1 to 6 for different demographic sub-samples. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

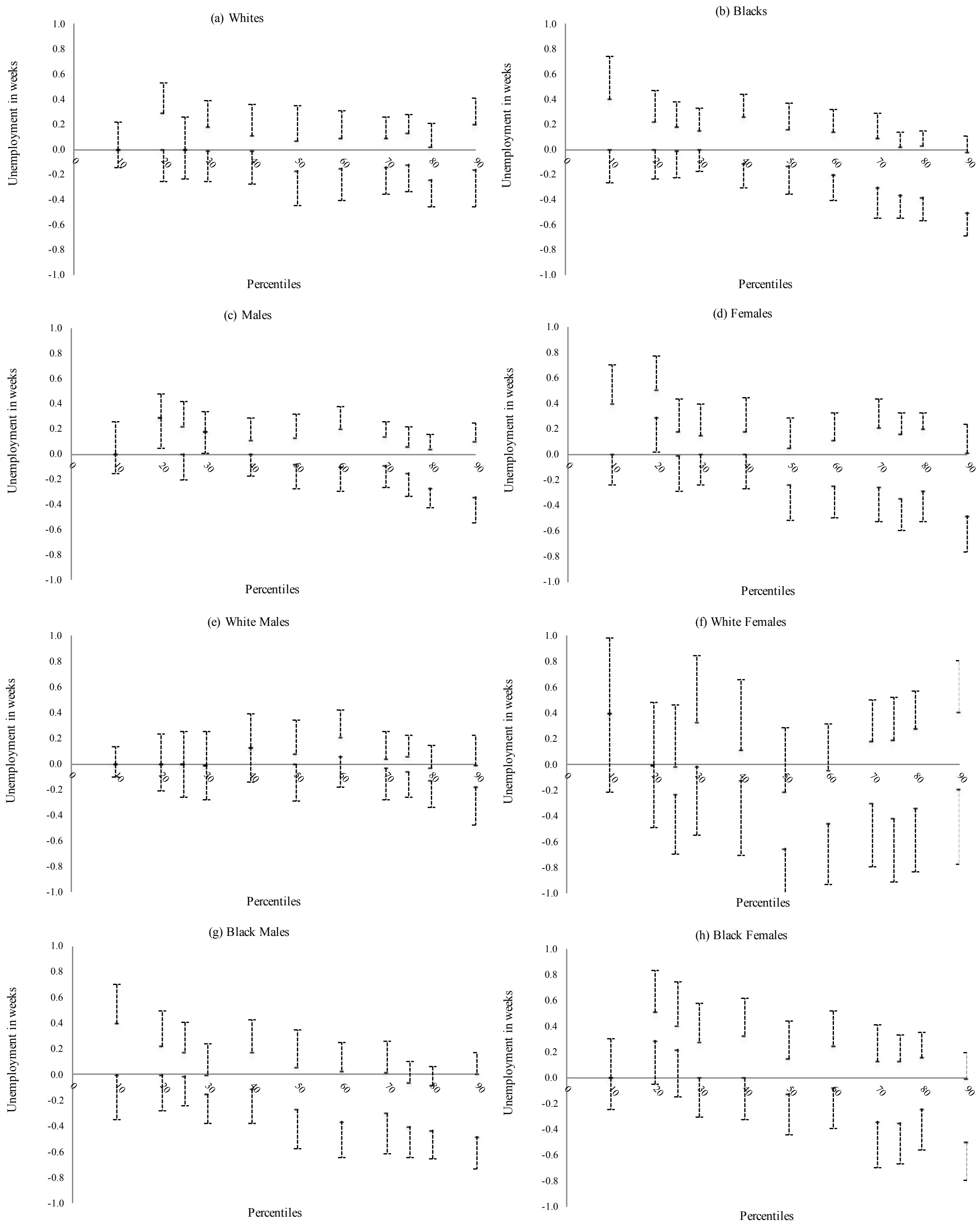

Figure 5. Bounds for CQTE on the log of Unemployment spells under Assumptions 1 to 7 for different demographic sub-samples. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

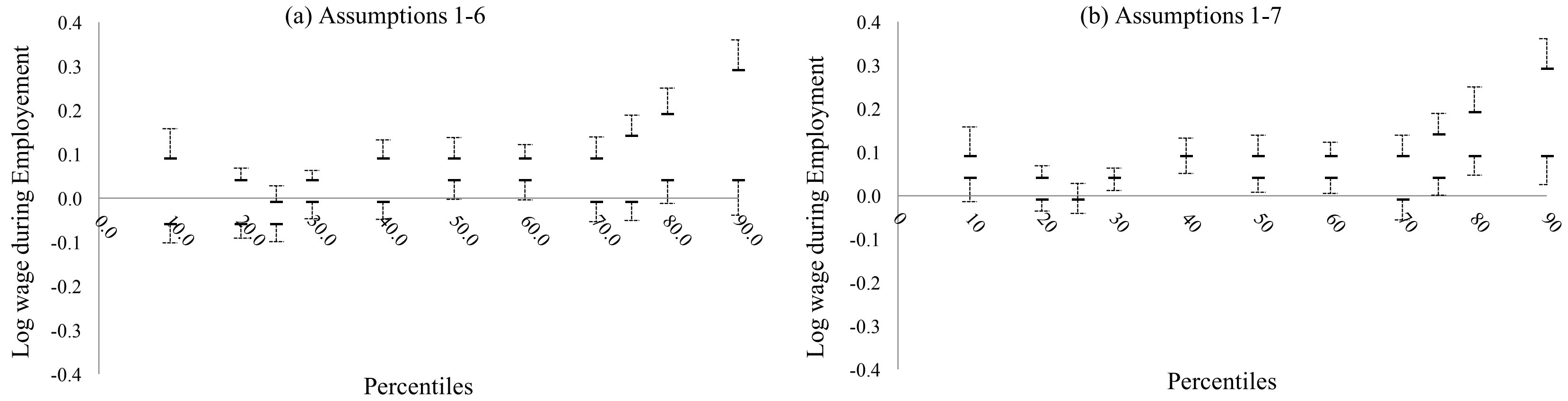

Figure 6. Bounds for CQTE on log wages during Employment spells for the Non-Hispanic sample. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.

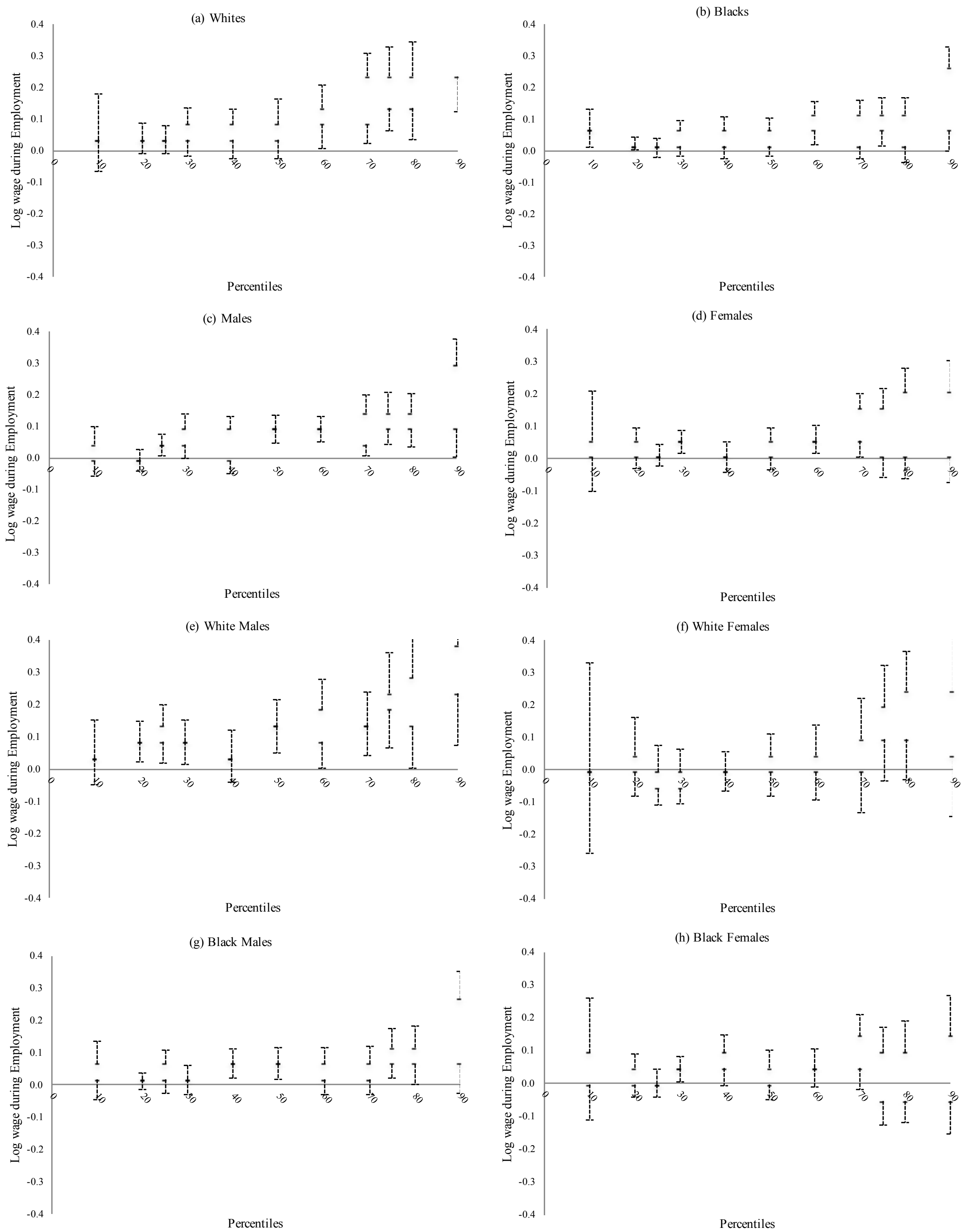

Figure 7. Bounds for CQTE on log wages during Employment spells under Assumptions 1 to 7 for different demographic sub-samples. Upper and lower bounds are denoted by a short dash, while the 90 percent Imbens and Manski (2004) confidence intervals are denoted by a long dash at the end of the dashed vertical line.