

TEMPLE
UNIVERSITY®

**Job Displacement during the Great Recession: Tight Bounds
on Distributional Treatment Effect Parameters using Panel
Data**

**Brantly Callaway
Department of Economics
Temple University**

Department of Economics
DETU Working Paper 17-03
December 2017

1301 Cecil B. Moore Avenue, Philadelphia, PA 19122

<http://www.cla.temple.edu/economics/faculty/detu-working-paper-series/>

Job Displacement during the Great Recession: Tight Bounds on Distributional Treatment Effect Parameters using Panel Data

Brantly Callaway^{*†}

This Version: August 2017
First Version: October 2015

Abstract

Late prime-age workers who were displaced during the Great Recession lost on average 39% of their earnings relative to their counterfactual earnings had they not been displaced. But the average effect masks substantial heterogeneity across workers. This paper develops new techniques to bound distributional treatment effect parameters that depend on the joint distribution of potential outcomes – an object not identified by standard identifying assumptions such as selection on observables or even when treatment is randomly assigned. I show that panel data and an additional assumption on the dependence between untreated potential outcomes for the treated group over time (i) provide more identifying power for distributional treatment effect parameters than existing bounds and (ii) provide a more plausible set of conditions than existing methods that obtain point identification.

Keywords: Joint Distribution of Potential Outcomes, Distribution of the Treatment Effect, Quantile of the Treatment Effect, Copula Stability Assumption, Panel Data, Job Displacement

JEL Codes: C14, C31, C33, J63

^{*}Department of Economics, Temple University, 1301 Cecil B. Moore Avenue, Ritter Annex 867, Philadelphia, PA 19122.
Email: brantly.callaway@temple.edu

[†]I would like to thank Alper Arslan, Colin Cameron, Bill Collins, Scott Cunningham, Dakshina De Silva, Juan Carlos Escanciano, Dalia Ghanem, Collin Green, Federico Gutierrez, Jorgen Hansen, Atsushi Inoue, Catherine Maclean, Pedro Sant’Anna, Shu Shen, Artyom Shneyerov, Doug Webber, and seminar participants at Baylor University, the Bureau of Labor Statistics, Concordia University, the Consumer Financial Protection Bureau, Indiana University, Lancaster University, Mathematica Policy Research, Temple University, the University of California-Davis, and the University of California-Merced for helpful comments. I especially thank my advisor Tong Li for his detailed comments and mentorship on this project.

1 Introduction

From the official beginning of the Great Recession in December 2007 to October 2009, the U.S. economy shed 8.4 million jobs and the unemployment rate doubled from 5.0% to 10.0%. Reemployment has been slow; many workers exited the labor force, had long unemployment spells, or moved into part time employment (Farber 2015). Using standard techniques, I find that annual earnings of displaced workers were, on average, 39% lower in 2011 than they would have been had the worker not been displaced. However, the effect of job displacement may be quite heterogeneous across workers. Some workers may quickly move into similar jobs, some may move into jobs with lower wages or into part time employment, and others may remain unemployed. Understanding this heterogeneity is of interest to researchers and policymakers. The policy response may be quite different if the effect of job displacement is very similar for all individuals compared to the case with very heterogeneous effects.

To understand the heterogeneous effects of job displacement, I develop new tight bounds on distributional treatment effect parameters that exploit having access to panel data. These bounds are much tighter than existing bounds and provide a credible alternative to point identifying assumptions that are not likely to hold in the current application. I find that workers in the 95th percentile of earnings losses due to displacement lose between 86% and 97% of earnings relative to counterfactual earnings had they not been displaced. I also find that at least 7.0% of workers have higher earnings after displacement than they would have had if they had not been displaced. These findings indicate that there is substantial heterogeneity in the effect of job displacement, but they would not be available using standard approaches to program evaluation. Finally, I find that, on average, earnings losses for workers with higher earnings before the recession are larger in magnitude than earnings losses for workers with lower earnings before the recession; but, as a fraction of earnings, average earnings losses are roughly constant across the distribution of pre-recession earnings.

Learning about heterogeneity in the effect of job displacement requires knowledge of the joint distribution of displaced potential earnings and non-displaced potential earnings for the group of workers that are displaced. However, the joint distribution of potential outcomes is not identified under common identifying assumptions such as selection on observables or even when individuals are randomly assigned to treatment. In each of these cases, although the marginal distributions of displaced and non-displaced potential earnings are identified, the copula – which “couples” the marginal distributions into the joint distribution and captures the dependence between the marginal distributions – is not identified.

To give an example, suppose a researcher is interested in the fraction of workers who have higher earnings following displacement than they would have had if they not been displaced. Further, suppose hypothetically that workers are randomly assigned to being displaced or not being displaced. In this case, the average effect of job displacement is identified – it is given by the difference between average earnings of those who are randomly assigned to be displaced and

those who are randomly assigned to not be displaced. But the fraction of workers that benefit from displacement is not identified because, for workers randomly assigned to be displaced (non-displaced), where they *would* be in the distribution of non-displaced (displaced) earnings is not known.

There are many important parameters that depend on the joint distribution of potential outcomes. These include the fraction of individuals that benefit from being treated, the correlation between treated and untreated potential outcomes, the variance of the treatment effect, the quantiles and distribution of the treatment effect itself, and the distribution of the treatment effect conditional on an individual's untreated potential outcome (Heckman, Smith, and Clements 1997; Firpo and Ridder 2008). These parameters are not just of theoretical importance. Policymakers may decide to implement a policy based on whether or not a large enough fraction of the population benefits from the policy rather than based on the average benefit of the policy. As another example, policymakers are likely to prefer programs with widespread though smaller benefits to ones where very few people benefit but have extremely large benefits. Finally, for some treatments that are not the direct result of policy decisions, such as job displacement, the response from policymakers may differ in cases where many individuals all experience a small effect of treatment compared to one where a few individuals are very affected.

Existing methods take two polar approaches to identifying the joint distribution of potential outcomes. One idea is to construct bounds on the joint distribution without imposing any assumptions on the unknown dependence (Heckman, Smith, and Clements 1997; Fan and Park 2009; Fan and Park 2012). In the case of job displacement, these bounds are not very informative. The main implications of these bounds are that (i) at least 22% of workers have lower earnings due to displacement than they would have had if they had not been displaced and (ii) the median of the treatment effect is between 65% lower earnings and 56% higher earnings.

Another approach is to assume that the dependence is known. The leading choice is perfect positive dependence.¹ This assumption says that individuals at a given rank in the distribution of earnings following displacement would have the same rank in the distribution of non-displaced potential earnings. This is a very strong assumption as it imposes severe restrictions on how heterogeneous the effect of job displacement can be; for example, it prohibits any workers at the top of the distribution of non-displaced earnings from becoming unemployed or taking a part time job following displacement. But the assumption is much stronger than that – it prohibits displacement from even swapping the rank of any workers relative to their rank had they not been displaced.

In light of (i) the implausibility of existing point-identifying assumptions and (ii) the wide bounds resulting from imposing no assumptions on the missing dependence, I develop new, tighter

¹This assumption was first implicitly made in the earliest work on estimating the distributional effects of treatment (Doksum 1974; Lehmann 1974) that compared the difference between treated quantiles and untreated quantiles and interpreted this difference as the treatment effect at that quantile. There is also recent work on testing the assumption of perfect positive dependence (Bitler, Gelbach, and Hoynes 2006; Dong and Shen 2015; Frandsen and Lefgren 2015)

bounds on parameters that depend on the joint distribution of potential outcomes. Unlike existing work which considers the case of cross-sectional data, I exploit having access to panel data on workers’ annual earnings. Panel data presents a unique opportunity to observe, at least for some individuals, both their displaced and non-displaced potential earnings though these are observed at different points in time. With panel data and under plausible identifying assumptions, the bounds on the joint distribution are much tighter – in theory, the joint distribution could even be point identified. To implement my method requires at least three periods of panel data.

Even though panel data appears to be useful for identifying the joint distribution of potential outcomes, there are still some challenges. In the context of Difference in Differences models, previous work has used panel data to recover missing dependence in the current period from observed dependence in previous periods (Callaway and Li 2017). But that approach is not possible in the current context because the dependence between displaced and non-displaced potential earnings is never observed – even in previous periods. Instead panel data is informative about the dependence between non-displaced potential earnings over time. One idea would be to assume perfect positive dependence between non-displaced potential earnings over time (Heckman and Smith 1998). This assumption results in point identification. With three periods of panel data, a researcher could pre-test this assumption by checking whether or not perfect positive dependence occurs in the periods before displacement. This assumption is rejected in the current application; intuitively, it requires no changes in ranks of annual earnings over time which is a very strong assumption.

Instead of assuming perfect positive dependence of non-displaced potential earnings over time, I assume that the dependence (or copula) of non-displaced potential earnings over time is the same over time. I call this assumption the Copula Stability Assumption. Recent work on income mobility decomposes the income at two different points in time into the marginal distributions – which capture inequality – and the copula which captures income mobility (Chetty, Hendren, Kline, and Saez 2014). Thus, in the context of job displacement, the Copula Stability Assumption requires that, in the absence of job displacement, earnings mobility would be constant over time for the group of displaced workers. Requiring that earnings mobility is constant over time is a considerably weaker assumption than assuming there is no earnings mobility over time. Importantly, the Copula Stability Assumption does not restrict the distribution of earnings over time. For example, the distribution of earnings can shift to the right over time or the distribution of earnings can become increasingly unequal over time.

I provide two pieces of evidence in favor of the Copula Stability Assumption. First, I show that the Copula Stability Assumption is likely to be satisfied in a very general model of the type typically estimated in panel data settings. Second, there is empirical evidence in favor of the Copula Stability Assumption. In the United States, despite large increases in inequality, there has been remarkably little change in yearly earnings mobility since the middle of the 20th century (Kopczuk, Saez, and Song 2010, and also see Figure 1).

The final requirement for using my method is that the counterfactual distribution of non-displaced potential earnings for the group of displaced workers must be identified. Because job displacement is not randomly assigned, this requires some type of identifying assumption. I use the Change in Changes model (Athey and Imbens 2006) to identify this distribution though the results are not sensitive to using other methods such as selection on observables (Firpo 2007) as long as a lag of earnings is included as a conditioning variable.

The bounds work in the following way. Let Y_{1t} be displaced potential earnings after displacement, Y_{0t} be non-displaced potential earnings after displacement, and Y_{0t-1} be observed non-displaced earnings before displacement. Existing bounds come from statistical bounds on bivariate distributions when the marginal distributions are known (Hoeffding 1940; Fréchet 1951). Under the setup in the current paper, the joint distributions of (Y_{1t}, Y_{0t-1}) and (Y_{0t}, Y_{0t-1}) are also available. I utilize the following result: for three random variables, when two of the three bivariate joint distributions are known, then bounds on the third bivariate joint distribution are at least as tight as the bounds when only the marginal distributions are known (Joe 1997).

Consider an extreme example. Suppose Y_{1t} and Y_{0t-1} are perfectly positively dependent and Y_{0t} and Y_{0t-1} are perfectly positively dependent, then Y_{1t} and Y_{0t} must also be perfectly positively dependent. In this case, the extra information from panel data results in point identification. In fact, point identification will occur when either (A) perfect positive dependence is observed between Y_{1t} and Y_{0t-1} or (B) perfect positive dependence is observed between Y_{0t-1} and Y_{0t-2} . The first case is very similar to the leading idea for point identification – perfect positive dependence across treated and untreated potential outcomes – though it also involves an additional time dimension. The second case is exactly the leading assumption for point identification with panel data – perfect positive dependence of untreated potential outcomes over time. Moreover, the bounds are tighter as either (Y_{1t}, Y_{0t-1}) or (Y_{0t}, Y_{0t-1}) becomes more positively dependent. This implies that even when these assumptions are violated, if these assumptions are “close” to holding, my method is robust to these deviations and will deliver tight bounds in precisely this case. Job displacement falls exactly into this category. Neither type of perfect positive dependence is observed; nonetheless, there is strong positive dependence which results in substantially tighter bounds.

Under the current setup, I am also able to study a parameter I call the Average Treatment Effect on the Treated Conditional on Previous Outcome (ATT-CPO). Although this parameter could be identified under some existing assumptions (for example, an experiment where panel data is also available), it is not under the standard panel data models used in much of the job displacement literature. I find that, on average, earnings losses for workers with higher earnings before the recession are larger in magnitude than earnings losses for workers with lower earnings before the recession; but, as a fraction of earnings, average earnings losses are very similar across the distribution of pre-recession earnings.

There are three other approaches to bounding the joint distribution of potential outcomes that

should be mentioned. Fan, Guerre, and Zhu (2017) bound parameters that depend on the joint distribution when covariates are available. This approach could theoretically be combined with the approach considered in the current paper to obtain even tighter bounds at the cost of significantly more challenging estimation. Another assumption that can bound parameters that depend on the joint distribution of potential outcomes is the assumption of Monotone Treatment Response (MTR) (Manski 1997). Kim (2014) combines this assumption with the statistical bounds approach. MTR would imply that earnings for displaced workers cannot be larger than earnings would have been had they not been displaced. This assumption is rejected by the bounds developed in the current paper. Finally, Frandsen and Lefgren (2016) obtain tighter bounds on the joint distribution of potential outcomes by ruling out negative dependence between the potential outcomes.

There is some empirical work studying the distributional effects of participating in a program. Djebbari and Smith (2008) use Fréchet-Hoeffding bounds to study the distributional effects of the PROGRESA program in Mexico. Carneiro, Hansen, and Heckman (2003) and Abbring and Heckman (2007), among others, use factor models to identify the joint distribution of treated and untreated potential outcomes.

2 Background

This section provides some context, motivation, and required details for studying distributional treatment effect parameters. After introducing some notation, it considers distributional treatment effect parameters that depend on the joint distribution of potential outcomes and why they are useful. Finally, it discusses in more detail why the joint distribution of potential outcomes is not identified under conventional identifying assumptions as well as existing stronger assumptions that point identify the joint distribution.

2.1 Treatment Effects Setup

The notation used throughout the paper is very similar to the notation used in the treatment effects literature in statistics and econometrics. All individuals in the population either participate or do not participate in a treatment. Let $D_t = 1$ for individuals that participate in the treatment and $D_t = 0$ for individuals who do not participate in the treatment (to minimize notation, a subscript i representing each individual is omitted). The paper considers the case where panel data is available and therefore random variables have a time subscript t . Each individual has potential outcomes in the treated and untreated states at time t which are given by Y_{1t} and Y_{0t} , respectively. But, for each individual, only one of these potential outcomes is observed at each time period. For individuals that are treated in period t , Y_{1t} is observed, but Y_{0t} is not observed. For individuals that are untreated in period t , Y_{0t} is observed but Y_{1t} is unobserved. Let Y_t be the

observed outcome in period t ; one can then write

$$Y_t = D_t Y_{1t} + (1 - D_t) Y_{0t}$$

It is helpful to make the following assumption

Assumption 1. *The observed data consists of $\{Y_{idt}, Y_{0it-1}, Y_{0it-2}, X_i, D_{it}\}$ and is independently and identically distributed.*

Assumption 1 covers available data in the baseline case considered in the paper. In particular, Assumption 1 says that the researcher observes outcomes in three periods. The researcher may also observe a vector of covariates X which, for simplicity, I assume are time invariant. Assumption 1 also says that individuals are first treated in the last period which implies that untreated potential outcomes are observed for both the treated group and the untreated group in periods $t - 1$ and $t - 2$. That is,

$$Y_{t-1} = Y_{0t-1} \text{ and } Y_{t-2} = Y_{0t-2}$$

Assumption 1 can be relaxed if additional periods are available or if treatment can occur in other periods besides the last one, but it represents a baseline case for tighter bounds and corresponds to the data used to study job displacement.

The next assumption is the starting point for the main identification results in the paper.

Assumption 2. $F_{Y_{1t}|D_t=1}$ and $F_{Y_{0t}|D_t=1}$ are identified.

Assumption 2 says that the marginal distribution of treated potential outcomes for the treated group, $F_{Y_{1t}|D_t=1}$, and the marginal distribution of untreated potential outcomes for the treated group, $F_{Y_{0t}|D_t=1}$, are identified. The first is not a strong assumption – it is given by the distribution of observed outcomes for the treated group, $F_{Y_t|D_t=1}$. The second is a stronger assumption. The counterfactual distribution would be identified if, for example, treatment were randomly assigned. However, in cases with observational data, like job displacement, it requires some identifying assumption. But there are many methods available to identify this counterfactual distribution. At any rate, the goal of the current paper is to go beyond the more standard objective of identifying this counterfactual distribution and learn about the joint distribution; thus, at this point, Assumption 2 considers the more standard problem of identifying the counterfactual marginal distribution to be solved.

2.2 Parameters of Interest

Assumption 2 implies that parameters that depend on the marginal distributions of treated and untreated potential outcomes for the treated group are identified. These include the Average

Treatment Effect on the Treated (ATT) ²

$$ATT = E[Y_{1t} - Y_{0t} | D_t = 1]$$

and the Quantile Treatment Effect on the Treated (QTT)

$$QTT(\tau) = F_{Y_{1t}|D_t=1}^{-1}(\tau) - F_{Y_{0t}|D_t=1}^{-1}(\tau)$$

for $\tau \in (0, 1)$ and where $F_X^{-1}(\tau) = \inf\{x : F_X(x) \geq \tau\}$.³

But distributional parameters that depend on the joint distribution of potential outcomes are not identified and these may be of considerable interest. For job displacement, I focus primarily on the Distribution of the Treatment Effect for the Treated (DoTT) and closely related parameters that are simple functionals of the DoTT; Heckman, Smith, and Clements (1997) and Firpo and Ridder (2008) discuss other parameters in this class that may be of interest in other applications.

The DoTT is the fraction of individuals that experience a treatment effect less than some value δ and is given by

$$DoTT(\delta) = P(Y_{1t} - Y_{0t} \leq \delta | D_t = 1)$$

One can estimate the DoTT for different values of δ and plot them. An alternative approach, and the one that seems more useful for studying job displacement is to invert the DoTT to obtain Quantiles of the Treatment Effect on the Treated (QoTT) which is given by

$$QoTT(\tau) = \inf\{\delta : DoTT(\delta) \geq \tau\}$$

To give some examples, in the context of job displacement, $QoTT(0.05)$ is the 5th percentile of the individual level effect of job displacement – these are the workers who experience some of the largest negative effects of job displacement. $QoTT(0.5)$ is the median effect of job displacement. And $QoTT(0.95)$ are workers who have close to the highest earnings relative to what they would have if they had not been displaced. Also, for some $0 < \tau_L < \tau_U < 1$, I use $QoTT(\tau_U) - QoTT(\tau_L)$ as a measure of treatment effect heterogeneity. An example is the difference between the 95th percentile of earnings following job displacement and the 5th percentile of earnings following job displacement.

²All the parameters mentioned above condition on being part of the treated group, but one may also be interested in these parameters for the entire population. Panel data is most useful for identifying parameters conditional on being part of the treated group. Using the techniques presented in the current paper can still lead to bounds on parameters for the entire population by combining the bounds for the treated group presented in the current paper with bounds for the untreated group coming from existing statistical bounds. These bounds will be tighter if a larger fraction of the population is treated. I do not pursue bounds on parameters for the entire population throughout the rest of the paper.

³Quantile Treatment Effects are very useful for evaluating welfare effects of a policy because they allow a researcher to compare the distribution of outcomes under the policy to the distribution of outcomes without the policy which provides considerably more information than just the knowledge of the average effect of the policy (Sen 1997; Carneiro, Hansen, and Heckman 2001).

Another interesting parameter for job displacement is the fraction of workers that have higher earnings following job displacement than they would have had if they had not been displaced. Let β denote this probability; it is given by

$$\beta = 1 - DoTT(0)$$

One important limitation of the QTT and the QoTT is that they can fail to provide information on which types of individuals experience the largest effects of treatment. For example, with panel data one can address whether workers with high earnings or low earnings in the previous period experience larger effects of job displacement. This parameter, which I term the Average Treatment Effect on the Treated Conditional on Previous Outcomes (ATT-CPO)⁴ is given by

$$ATT-CPO(y') = E[Y_{1t} - Y_{0t} | Y_{0t-1} = y', D_t = 1]$$

In the current framework, this parameter is point identified because the joint distributions $F_{Y_{1t}, Y_{0t-1} | D_t=1}$ and $F_{Y_{0t}, Y_{0t-1} | D_t=1}$ are identified, and the ATT-CPO depends only on these joint distributions and not on the joint distribution of treated and untreated potential outcomes for the treated group in period t . The next example provides a case where a combination of the ATT-CPO and the QoTT can provide a better understanding of the distributional effect of treatment compared to the QTT.

Example 1. *Suppose there are 10 workers in the population that are observed in two periods. In the first period, suppose each worker's earnings are given by their number; worker 1 has earnings of 1, worker 2 has earnings of 2, etc. In the second period, in the absence of displacement, suppose each worker keeps the same earnings as in the first period. For worker 10, suppose his earnings decrease to 0 if he is displaced, but for the other workers, suppose they are able to find a new job with the same earnings. In this case, the QTT is constant everywhere and equal to -1. In much applied research, this effect would wrongly be interpreted as the effect of displacement being the same across workers with high and low earnings. On the other hand, $QoTT(0.1) = -10$ and $QoTT(\tau) = 0$ for $\tau > 0.1$. Immediately, this would imply that the effect of treatment is very heterogeneous – one worker has much lower earnings due to being displaced while most workers experience no effect of displacement on earnings. Also, $ATT-CPO(10) = -10$ but $ATT-CPO(y') = 0$ for other values of y' . This would imply the effect of displacement is strongest for workers with highest earnings.*

⁴Heckman, Smith, and Clements (1997) suggest the related parameter $F_{Y_{1t}-Y_{0t} | Y_{0t}}(\delta | y_0)$. This is the distribution of the treatment effect conditional on the base state Y_{0t} taking some particular value y_0 . It is an interesting parameter, but it suffers from being difficult to display graphically because in most cases a researcher is interested in this parameter while varying y' in many values of its support. A plot of the result would be a three dimensional and difficult to interpret. An alternative measure is $E[Y_{1t} - Y_{0t} | Y_{0t} = y_0]$. This is the average treatment effect conditional on the base state Y_{0t} taking some particular value y_0 . Varying y_0 results in an easy to interpret two dimensional plot. However, in the current setup, this parameter is not point identified because it depends on the joint distribution of treated and untreated potential outcomes which is only partially identified in the current setup.

This example is less extreme than it appears. If any workers from the middle or top of the non-displaced potential earnings distribution move to the lower part of the distribution of earnings following displacement – which could happen due to difficulty finding new employment, moving to part time work, or having much lower wages at a new job – then the QTT will be very difficult to interpret. However, the QoTT and the ATT-CPO can still be very helpful to understand the distributional effects of displacement.

2.3 The Identification Issue and Existing Solutions

This section explains in greater detail the fundamental reason why the joint distribution of potential outcomes is not point identified except under strong assumptions. First, I assume that both the marginal distribution of treated potential outcomes for the treated group $F_{Y_{1t}|D_t=1}$ and the marginal distribution of untreated potential outcomes for the treated group $F_{Y_{0t}|D_t=1}$ are identified. The first can be obtained directly from the data; the second is obtained under some identifying assumption which is assumed to be available. Sklar (1959) demonstrates that joint distributions can be written as the copula function of marginal distributions in the following way

$$F_{Y_{1t}, Y_{0t}|D_t=1}(y_1, y_0) = C_{Y_{1t}, Y_{0t}|D_t=1} \left(F_{Y_{1t}|D_t=1}(y_1), F_{Y_{0t}|D_t=1}(y_0) \right) \quad (2.1)$$

where $C_{Y_{1t}, Y_{0t}|D_t=1}(\cdot, \cdot) : [0, 1]^2 \rightarrow [0, 1]$. This representation highlights the key piece of missing information under standard assumptions – the copula function. Using results from the statistics literature, one can still construct the so-called Fréchet-Hoeffding bounds on the joint distribution. These bounds arise from considering two extreme cases: (i) when there is perfect positive dependence between the two marginal distributions and (ii) when there is perfect negative dependence between the two distributions. Heckman, Smith, and Clements (1997) follow this procedure and find that it leads to very wide bounds in general. Moreover, that paper points out that under strong forms of negative dependence, the bounds do not seem to make sense in an application on the treatment effect of participating in a job training program.

At the other extreme, one could posit a guess for the copula. In the cross-sectional case, the most common assumption is perfect positive dependence between treated potential outcomes and untreated potential outcomes for the treated group. This assumption can be written as

Assumption 3 (CS-PPD).

$$F_{Y_{1t}|D_t=1}(Y_{1t}) = F_{Y_{0t}|D_t=1}(Y_{0t})$$

Assumption 3 implies that

$$Y_{0t} = F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{1t}|D_t=1}(Y_{1t}))$$

which means that for any individual in the treated group with observed outcome Y_{1t} , their coun-

terfactual untreated potential outcome Y_{0t} is also known which implies that the joint distribution is point identified. Although this assumption might be more plausible than assuming independence or perfect negative dependence, it seems very unlikely to hold in practice because it severely restricts the ability of treatment to have different effects across different individuals. But the idea that different individuals can experience different effects of treatment is one of the central themes of the entire treatment effects literature. In the context of job displacement, perfect positive dependence seems unlikely to hold because it would prohibit individuals at the top of the pre-displacement earnings distribution from being unemployed or taking a part time job following job displacement.

With panel data, perhaps a more plausible assumption is perfect positive dependence in untreated potential outcomes over time (this idea is mentioned in Heckman and Smith 1998):

Assumption 4 (Panel-PPD).

$$F_{Y_{0t}|D_t=1}(Y_{0t}) = F_{Y_{0t-1}|D_t=1}(Y_{0t-1})$$

Assumption 4 does not directly replace the unknown copula in Equation 2.1, but the next result establishes that this assumption also leads to point identification of the joint distribution of potential outcomes.

Result 1. *Under perfect positive dependence between untreated potential outcomes for the treated group over time,*

$$F_{Y_{1t}, Y_{0t}|D_t=1}(y_1, y_0) = F_{Y_{1t}, Y_{0t-1}|D_t=1}(y_1, F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0)))$$

When the researcher has access to more than two periods of panel data, one can apply a sort of pre-test to this assumption. That is, one can check whether perfect positive dependence in untreated potential outcomes holds between periods $t - 1$ and $t - 2$ and this can provide evidence as to whether or not perfect positive dependence is likely to hold between periods t and $t - 1$. In the application in the current paper, I find that this assumption does not hold, but it is not too far from holding; in other words, in the absence of being displaced, workers do change ranks in the distribution of earnings over time, but, for the most part, the change in ranks is small.

3 Identification

In the previous section, I have argued that assumptions that directly replace the unknown copula in Equation 2.1 are not likely to hold. This section considers an alternative approach that does not substitute for the copula in Equation 2.1 directly but limits the possibilities for the copula.

The next assumption is the main identifying assumption in the paper.

Copula Stability Assumption. For all $(u, v) \in [0, 1]^2$

$$C_{Y_{0t}, Y_{0t-1}|D_t=1}(u, v) = C_{Y_{0t-1}, Y_{0t-2}|D_t=1}(u, v)$$

The Copula Stability Assumption says that the dependence between untreated potential outcomes at periods t and $t - 1$ is the same as the dependence between untreated potential outcomes at periods $t - 1$ and $t - 2$. This assumption is useful because the dependence between untreated potential outcomes at period t and period $t - 1$ is not observed. Although, by assumption, the counterfactual distribution of untreated potential outcomes for the treated group, $F_{Y_{0t}|D_t=1}$, is identified and the distribution of untreated potential outcomes for the treated at period $t - 1$, $F_{Y_{0t-1}|D_t=1}$, is identified because untreated potential outcomes are observed for the treated group at period $t - 1$, their dependence is not identified because Y_{0t} and Y_{0t-1} are not simultaneously observed for the treated group. The Copula Stability Assumption recovers the missing dependence. This implies that the joint distribution of untreated potential outcomes at times t and $t - 1$ for the treated group, $F_{Y_{0t}, Y_{0t-1}|D_t=1}$, is identified. This joint distribution is not of primary interest in the current paper. But knowledge of this joint distribution is important for deriving tight bounds on the distributions and parameters of interest.

To better understand the Copula Stability Assumption, it is helpful to consider some examples. As a first example, the Copula Stability Assumption says that if untreated potential outcomes at period $t - 1$ are independent (or perfectly positively dependent) of untreated potential at period $t - 2$, then untreated outcomes at period t will continue to be independent (or perfectly positively dependent) of untreated outcomes at period $t - 1$. Or, for example, suppose the copula for $(Y_{0t-1}, Y_{0t-2}|D_t = 1)$ is Gaussian with parameter ρ , the Copula Stability Assumption says that the copula for $(Y_{0t}, Y_{0t-1}|D_t = 1)$ is also Gaussian with parameter ρ though the marginal distributions of outcomes can change in unrestricted ways. For example, the distribution of earnings can shift over time or could become more unequal over time. Likewise, if the copula is Archimedean, the Copula Stability Assumption says that the generator function does not change over time. For Archimedean copulas with a scalar parameter having a one-to-one mapping to dependence parameters such as Kendall's Tau or Spearman's Rho (examples include common Archimedean copulas such as the Clayton, Frank, and Gumbel copulas), the Copula Stability Assumption says that the dependence parameter is the same over time.⁵

Assumption 5. (*Outcomes are continuously distributed*)

Y_t, Y_{0t-1} and Y_{0t-2} are continuously distributed for $D_t = 0, 1$.

Assumption 5 allows for the quantile functions in the expressions below to be well-defined. In

⁵One could also make the Copula Stability Assumption conditional on some covariates X . This type of assumption might be more plausible in some applications. For example, earnings over time may be more strongly positively dependent for older workers than for younger workers. It should be noted, however, that the unconditional Copula Stability Assumption does not preclude covariates affecting outcomes; but it does place some restrictions on how covariates can affect the outcome of interest and especially how the effect of covariates changes over time. One cost of this approach is that nonparametric estimation would be significantly more challenging or strong parametric assumptions might be required to estimate the model.

the application on job displacement, the primary restriction of the Assumption 5 is that observations with 0 earnings need to be dropped, but this is standard practice in the job displacement literature (Jacobson, LaLonde, and Sullivan 1993; Stevens 1997; Kletzer and Fairlie 2003; Von Wachter, Song, and Manchester 2009) and is the approach followed in the current paper as well.

Lemma 1. *Under Assumptions 1, 2 and 5 and the Copula Stability Assumption,*

$$F_{Y_{0t}, Y_{0t-1}|D_t=1}(y_0, y') = F_{Y_{0t-1}, Y_{0t-2}|D_t=1}\left(F_{Y_{0t-1}|D_t=1}^{-1} \circ F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-2}|D_t=1}^{-1} \circ F_{Y_{0t-1}|D_t=1}(y')\right)$$

and

$$F_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y') = F_{Y_{0t-1}|Y_{0t-2}, D_t=1}\left(F_{Y_{0t-1}|D_t=1}^{-1} \circ F_{Y_{0t}|D_t=1}(y_0) \middle| F_{Y_{0t-2}|D_t=1}^{-1} \circ F_{Y_{0t-1}|D_t=1}(y')\right)$$

Lemma 1 shows that the joint distribution of Y_{0t} and Y_{0t-1} is identified under the Copula Stability Assumption. This joint distribution is an important building block in developing tighter bounds on the joint distribution of Y_{1t} and Y_{0t} for the treated group.

Next, I show how the Copula Stability Assumption can be used to derive tighter bounds on the joint distribution of potential outcomes. The next result is a simple application of Fréchet-Hoeffding bounds to a conditional distribution; it provides an important building block for constructing tighter bounds on the joint distribution of potential outcomes.

Lemma 2. *Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, bounds on the joint distribution of treated and untreated potential outcomes for the treated group conditional on outcomes in the previous period are given by*

$$F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^L(y_1, y_0|y') \leq F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}(y_1, y_0|y') \leq F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^U(y_1, y_0|y')$$

where

$$\begin{aligned} F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^L(y_1, y_0|y') &= \max\{F_{Y_{1t}|Y_{0t-1}, D_t=1}(y_1|y') + F_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y') - 1, 0\} \\ F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^U(y_1, y_0|y') &= \min\{F_{Y_{1t}|Y_{0t-1}, D_t=1}(y_1|y'), F_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y')\} \end{aligned}$$

The next theorem is the main result for bounds on the joint distribution of potential outcomes for the treated group.

Theorem 1. *Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, bounds on the joint distribution of treated and untreated potential outcomes for the treated group are given by*

$$F_{Y_{1t}, Y_{0t}|D_t=1}^L(y_1, y_0) \leq F_{Y_{1t}, Y_{0t}|D_t=1}(y_1, y_0) \leq F_{Y_{1t}, Y_{0t}|D_t=1}^U(y_1, y_0)$$

where

$$\begin{aligned} F_{Y_{1t}, Y_{0t}|D_t=1}^L(y_1, y_0) &= E[F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^L(y_1, y_0|Y_{0t-1})] \\ F_{Y_{1t}, Y_{0t}|D_t=1}^U(y_1, y_0) &= E[F_{Y_{1t}, Y_{0t}|Y_{0t-1}, D_t=1}^U(y_1, y_0|Y_{0t-1})] \end{aligned}$$

The bounds in Theorem 1 warrant some more discussion. First, these bounds will be tighter than the bounds without using panel data unless Y_{0t-1} is independent of Y_{1t} and Y_{0t} . But in most applications in economics Y_{0t} and Y_{0t-1} are likely to be positively dependent. On the other hand, the joint distribution will be point identified if either (i) Y_{1t} and Y_{0t-1} are perfectly positively dependent or (ii) Y_{0t} and Y_{0t-1} are perfectly positively dependent. Item (i) is very similar to the assumption of perfect positive dependence across treated and untreated groups (though it also includes a time dimension); Item (ii) is exactly the condition of perfect positive dependence in untreated potential outcomes over time used as a point identifying assumption (Heckman and Smith 1998). Together, these conditions imply that if either one of two natural limiting conditions hold in the data, then the joint distribution of potential outcomes will be point identified. Moreover, intuitively the bounds will be tighter in cases that are “closer” to either of these two limiting cases. This means that even in the case where the limiting conditions do not hold exactly, one is still able to (substantially) tighten the bounds that would arise in the case without panel data. I provide the intuition for this point in the next example and provide a more formal proof in the proposition that follows.

Example 2. *Spearman’s Rho is the correlation of the ranks of two random variables; i.e. $\rho_S = \text{Corr}(F_1(X_1), F_2(X_2))$. Bounds on Spearman’s Rho can be derived when two out of three joint distributions and all marginal distributions (exactly our case) are known (Joe 2015, Theorem 8.19). Because the marginal distributions $F_{Y_{1t}|D_t=1}(Y_{1t})$, $F_{Y_{0t}|D_t=1}(Y_{0t})$, and $F_{Y_{0t-1}|D_t=1}(Y_{0t-1})$ are uniformly distributed, their covariance matrix is given by*

$$\text{Cov} \left(F_{Y_{1t}|D_t=1}(Y_{1t}), F_{Y_{0t}|D_t=1}(Y_{0t}), F_{Y_{0t-1}|D_t=1}(Y_{0t-1}) \right) = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$$

Consider the case where ρ_{13} and ρ_{23} are identified and ρ_{12} is not known. ρ_{12} is partially identified because the covariance matrix must be positive semi-definite.

This results in the condition that

$$\rho_{13}\rho_{23} - \sqrt{\rho_{13}^2\rho_{23}^2 + (1 - \rho_{13}^2 - \rho_{23}^2)} \leq \rho_{12} \leq \rho_{13}\rho_{23} + \sqrt{\rho_{13}^2\rho_{23}^2 + (1 - \rho_{13}^2 - \rho_{23}^2)}$$

The width of the bounds is given by

$$\text{width} = 2\sqrt{\rho_{13}^2\rho_{23}^2 + (1 - \rho_{13}^2 - \rho_{23}^2)}$$

It is easy to show that for fixed ρ_{23} with $|\rho_{23}| < 1$, the width of the bounds on ρ_{12} are decreasing as ρ_{13} increases for $\rho_{13} > 0$, and width of the bounds are decreasing as ρ_{13} decreases for $\rho_{13} < 0$. When either ρ_{13} or ρ_{23} is equal to one in absolute value, ρ_{12} is point identified. This corresponds exactly to the case of perfect positive dependence (or perfect negative dependence) mentioned above for point identification. The intuition of this result is that as the copula moves “closer” to perfect positive dependence or perfect negative dependence, the bounds on the joint distribution of interest shrink.

Proposition 1. Fix the marginal distributions $F_{Y_{1t}|D_t=1}$, $F_{Y_{0t}|D_t=1}$, and $F_{Y_{0t-1}|D_t=1}$ and the conditional distribution $F_{Y_{1t}|Y_{0t-1}, D_t=1}$. Now consider two possibilities for $F_{Y_{0t}|Y_{0t-1}, D_t=1}$ given by F_1 and F_2 . Assume that F_1 , F_2 , and $F_{Y_{1t}|Y_{0t-1}, D_t=1}$ are stochastically increasing and that $F_{Y_{1t}|Y_{0t-1}, D_t=1} \prec^{SI} F_1 \prec^{SI} F_2$ where $F \prec^{SI} G$ indicates that G is more stochastically increasing than F ; more stochastically increasing is a common dependence ordering (see Joe (1997) and Nelsen (2007)). Then, the bounds on the joint distribution given in Theorem 1 are at least as tight when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_2$ as when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_1$.

Proposition 1 is a key result in the paper. It says that the bounds in the paper get tighter when there is stronger dependence between Y_{0t} and Y_{0t-1} (under the Copula Stability Assumption, this will be true when the dependence between Y_{0t-1} and Y_{0t-2} is stronger). An analogous result also holds for Y_{1t} and Y_{0t-1} – if the dependence is strong, the bounds will be tight. In the literature, assumptions of perfect positive dependence have been made as approximations because in many applications there is strong positive dependence though less than perfect positive dependence. Proposition 1 implies that the results in the current paper will be valid when perfect positive dependence assumptions are violated, but the bounds will be “tight” in the case where these assumptions are not too far from the truth. This is likely to be the most relevant case in many applications. When the bounds are applied to job displacement later, I show that there is strong positive dependence (both between Y_{1t} and Y_{0t-1} and between Y_{0t-1} and Y_{0t-2} for the treated group) but less than perfect positive dependence. This implies that the assumptions of perfect positive dependence will be violated, but it also implies that the bounds can be tightened substantially over bounds that only use the information from the marginal distributions of Y_{1t} and Y_{0t} , respectively.

Just as knowledge of $F_{Y_{1t}, Y_{0t-1}|D_t=1}$ and $F_{Y_{0t}, Y_{0t-1}|D_t=1}$ leads to bounds on the joint distribution of interest $F_{Y_{1t}, Y_{0t}|D_t=1}$, knowledge of these distributions can also be used to bound the DoTT, the QoTT, and other parameters that depend on the joint distribution. These results are presented next.

Sharp bounds on the distribution of the treatment effect are known in the case where there is

no additional information besides the marginal distributions (Fan and Park 2010). These bounds are obtained using results from the statistics literature for the distribution of the difference of two random variables when the marginal distributions are fixed (Makarov 1982; Rüschendorf 1982; Frank, Nelsen, and Schweizer 1987; Williamson and Downs 1990). I use these same bounds for the conditional joint distribution.

Lemma 3. *(Conditional Distribution of the Treatment Effect) Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, bounds on the distribution of the treatment effect for the treated group conditional on the outcome in the previous period are given by*

$$F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|y') \leq F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}(\delta|y') \leq F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|y')$$

where

$$\begin{aligned} F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|y') &= \sup_y \max\{F_{Y_{1t}|Y_{0t-1}, D_t=1}(y|y') - F_{Y_{0t}|Y_{0t-1}, D_t=1}(y - \delta|y'), 0\} \\ F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|y') &= 1 + \inf_y \min\{F_{Y_{1t}|Y_{0t-1}, D_t=1}(y|y') - F_{Y_{0t}|Y_{0t-1}, D_t=1}(y - \delta|y'), 0\} \end{aligned}$$

The next result provides bounds for the DoTT.

Theorem 2. *(Distribution of the Treatment Effect) Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, bounds on DoTT(δ) are given by*

$$DoTT^L(\delta) \leq DoTT(\delta) \leq DoTT^U(\delta)$$

where

$$\begin{aligned} DoTT^L(\delta) &= F_{Y_{1t}-Y_{0t}|D_t=1}^L(\delta) = E[F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|Y_{0t-1})|D_t = 1] \\ DoTT^U(\delta) &= F_{Y_{1t}-Y_{0t}|D_t=1}^U(\delta) = E[F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|Y_{0t-1})|D_t = 1] \end{aligned}$$

where $F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|y')$ and $F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|y')$ are given in Lemma 3.

Remark 1. *The results of Proposition 1 also imply that the bounds on the DoTT are tighter under the conditions given in Proposition 1. To be specific, consider two possibilities for $F_{Y_{0t}|Y_{0t-1}, D_t=1}$ given by F_1 and F_2 with $F_1 \prec^{SI} F_2$ as in Proposition 1 and assume all other conditions as in that proposition. Proposition 1 implies $C_1^L \prec^C C_2^L$ where C_j^L denotes the lower bound on the copula of Y_{1t} and Y_{0t} when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_j$ for $j = 1, 2$. Bounds on the DoTT get tighter when the lower bound of the copula is more concordant (Williamson and Downs 1990). This implies that the bounds on the DoTT will be tighter when there is more positive dependence between Y_{0t} and Y_{0t-1} (which occurs under the Copula Stability Assumption when there is more positive dependence between Y_{0t-1} and Y_{0t-2}).*

Bounds on the QoTT can be obtained from the bounds on the DoTT. The upper bound on the QoTT comes from inverting the lower bound of the DoTT, and the lower bound on the QoTT comes from inverting the upper bound on the DoTT.

Theorem 3. (*Quantile of the Treatment Effect*) Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, bounds on $QoTT(\tau)$ are given by

$$QoTT^L(\tau) \leq QoTT(\tau) \leq QoTT^U(\tau)$$

where

$$QoTT^L(\tau) = \inf\{\delta : DoTT^U(\delta) \geq \tau\}$$

$$QoTT^U(\tau) = \inf\{\delta : DoTT^L(\delta) \geq \tau\}$$

and $DoTT^L(\delta)$ and $DoTT^U(\delta)$ are given in Theorem 2.

Point Identification of the ATT-CPO Next, I show that the ATT-CPO is identified in the current setup. The reason why this parameter is point identified is that it depends on the joint distributions $F_{Y_{1t}, Y_{0t-1}|D_t=1}$ and $F_{Y_{0t}, Y_{0t-1}|D_t=1}$ which are both point identified, but it does not depend on the joint distribution of treated and untreated potential outcomes $F_{Y_{1t}, Y_{0t}|D_t=1}(y_1, y_0)$ which is only partially identified. The next result provides an explicit expression for this result that can be estimated using the observed data and the identified marginal distributions.

Theorem 4. Under Assumptions 1, 2 and 5 and the Copula Stability Assumption, the ATT-CPO is point identified and given by

$$\begin{aligned} ATT-CPO(y') &= E[Y_{1t} - Y_{0t}|Y_{0t-1} = y', D_t = 1] \\ &= E[Y_{1t}|Y_{0t-1} = y', D_t = 1] \\ &\quad - E[F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(Y_{0t-1})) \mid Y_{0t-2} = F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))] \end{aligned}$$

Theorem 4 shows that the ATT-CPO is identified and all of the terms in the theorem are fairly straightforward to estimate. The term $E[Y_{0t}|Y_{0t-1} = y', D_t = 1]$ is more challenging because the sampling process does not provide draws from the joint distribution $F_{Y_{0t}, Y_{0t-1}|D_t=1}$. Under the Copula Stability Assumption, one can use dependence information about Y_{0t-1} and Y_{0t-2} for the treated group – and the sampling process does provide draws from this distribution. However, the marginal distributions need to be adjusted which is the reason for the terms involving the quantile of the distribution in Theorem 4.

3.1 Empirical Evidence on the Copula Stability Assumption

This section provides some empirical evidence that the Copula Stability Assumption may be valid when the outcome of interest is yearly income – a leading case in labor economics. In this case, the Copula Stability Assumption says that income mobility, which has been interpreted as the copula of income over time in studies of mobility (Chetty, Hendren, Kline, and Saez 2014) or very similarly as the correlation between the ranks of income over time (Kopczuk, Saez, and Song 2010),⁶ is the same over time.⁷

A simple way to check if the copula is constant over time is to check if some dependence measure such as Spearman’s Rho or Kendall’s Tau is constant over time.⁸ Using administrative data from 1937-2003, Kopczuk, Saez, and Song (2010) find that the rank correlation (Spearman’s Rho) of yearly income is nearly constant in the U.S. Immediately following World War II, there was a slight decline in income mobility. Since then, there has been remarkable stability in income mobility (See Figure 1).

Moreover, Figure 1 also confirms the intuition that there is strong positive dependence of yearly income over time though the dependence is less than perfect positive dependence. This is precisely the case where the method developed in the current paper is likely to (i) provide more credible results than employing a perfect positive dependence over time assumption while (ii) yielding much tighter bounds on the joint distribution of potential outcomes than would be available using other methods that rely on purely statistical results to bound distributional treatment effects that depend on the joint distribution of potential outcomes.

3.2 Theoretical Evidence on the Copula Stability Assumption

In order to estimate the joint distribution of treated and untreated potential outcomes for the treated group, one must first identify and estimate the counterfactual distribution of untreated potential outcomes for the treated group $F_{Y_{0t}|D_t=1}$; that is, the distribution of earnings that displaced workers would have had if they had not been displaced must be identified. To identify this distribution, I use the Change in Changes (CIC) model (Athey and Imbens 2006). The CIC model is based on the following setup:

$$Y_{0is} = h_s(U_{is}) \quad \text{for } s = t, t - 1$$

⁶The dependence measure Spearman’s Rho is exactly the correlation of ranks. Dependence measures such as Spearman’s Rho or Kendall’s Tau are very closely related to copulas; for example, these dependence measures depend only on the copula of two random variables not the marginal distributions. Dependence measures also have the property of being ordered. For example, larger Spearman’s Rho indicates more positive dependence; two copulas, on the other hand, cannot generally be ordered. See Nelsen (2007) and Joe (2015) for more discussion on the relationship between dependence measures and copulas.

⁷It is also very similar to other work in the income mobility literature that considers transitions from one quintile of earnings in one period to another quintile of earnings in another period (Hungerford 1993; Gottschalk 1997; Carroll, Joulfaian, and Rider 2007)

⁸It is possible for a copula to change over time and have the same value of the dependence measure, but if the dependence measure changes over time, then the copula necessarily changes over time.

with (i) $h_s(u)$ strictly increasing in u and (ii) $U_s|D_t = d \sim F_{U|D_t=d}$ for $s = t, t-1$ and $d = 0, 1$. In addition to these conditions, I also assume $Y_{0it-2} = h_{t-2}(U_{it-2})$, $U_{is} = C_i + V_{is}$ for $s = t, t-1, t-2$, and $V_s|C, D_t = d \sim F_{V|C, D_t=d}$ for all $s = t, t-1, t-2$ (this last condition implies $U_s|D_t = d \sim F_{U|D_t=d}$ in each time period). These extra conditions simply extend the model to three periods and from the case with repeated cross sections to panel data. Importantly they allow the distribution of C to differ across the treated and untreated group. I call the set of conditions above the Three Period Panel CIC model.

Proposition 2. *In the Three Period Panel CIC model given above, and under the additional assumption that $F_{V_t, V_{t-1}|C, D_t=1} = F_{V_{t-1}, V_{t-2}|C, D_t=1}$, the Copula Stability Assumption holds.*

Proposition 2 gives the additional condition required for the Copula Stability Assumption to hold in the CIC model. The additional condition is weak and would be satisfied if the V_s are iid, but it can also allow for dependence of V_s over time as long as the dependence is constant.

4 Estimation

This section briefly discusses estimation and inference the QoTT and ATT-CPO under the identification results presented above and supposing that an estimate of the counterfactual distribution of untreated potential outcomes for the treated group, $\hat{F}_{Y_{0t}|D_t=1}$, is available.

Estimating Distributions, Quantiles, and Conditional Distributions

Both the QoTT and the ATT-CPO require estimating distribution and quantile functions. I estimate unconditional distributions by their empirical cdf:

$$\hat{F}_Z(z) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{Z_i \leq z\}$$

To estimate quantile functions, I invert empirical cdfs:

$$\hat{F}_Z^{-1}(\tau) = \inf\{z : \hat{F}_Z(z) \geq \tau\}$$

Estimating the QoTT also requires estimating conditional distribution functions. For estimating a conditional distribution function, I use local logit (Fan and Gijbels 1996; Fan, Farmen, and Gijbels 1998; Frölich 2006):

$$\hat{F}_{Z|Z'}(z|z') = \Lambda(\hat{\beta}_0(z, z') + \hat{\beta}_1(z, z')z')$$

where

$$(\hat{\beta}_0(z, z'), \hat{\beta}_1(z, z')) = \arg \max_{\beta_0(z, z'), \beta_1(z, z')} \sum_{i=1}^n \left(\mathbb{1}\{Z_i \leq z\} \log(\Lambda(\beta_0(z, z') + \beta_1(z, z')Z')) \right. \\ \left. + (1 - \mathbb{1}\{Z_i \leq z\}) \log(1 - \Lambda(\beta_0(z, z') + \beta_1(z, z')Z')) \right) K_h(Z'_i - z')$$

In other words, $\beta_0(z, z')$ and $\beta_1(z, z')$ are parameters indexed by z and z' . Here, h is a bandwidth satisfying $h \rightarrow 0$ as $n \rightarrow \infty$ and $nh \rightarrow \infty$; K is a kernel function. The local logit model guarantees that estimates of the conditional distributions are bounded between 0 and 1 (e.g. estimates using a local linear model would not guarantee this; see Hall, Wolff, and Yao (1999) and Hansen (2004)). However, it is not guaranteed to be increasing in z . Thus, I use rearrangement (Chernozhukov, Fernandez-Val, and Galichon 2010), which should have not have any effect on the asymptotic properties of the estimators and guarantees that the conditional distributions are invertible.

Estimating the QoTT

Estimation of the QoTT is based on the results of Lemma 1, Lemma 3, Theorem 2, and Theorem 3. Broadly speaking, estimation proceeds by (i) estimating $F_{Y_{1t}|Y_{0t-1}, D_t=1}$ and $F_{Y_{0t}|Y_{0t-1}, D_t=1}$, (ii) plugging these into the results in Lemma 3 to obtain bounds on the conditional distribution of the treatment effect, (iii) averaging the conditional bounds over the values of Y_{0t-1} for the treated group to obtain bounds on the DoTT, and (iv) inverting the bounds on the DoTT to obtain the QoTT. Although it takes several steps, the estimation is fairly straightforward. It is possible to use plug-in estimators for every term except $F_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y')$. This term is identified under the Copula Stability Assumption and Lemma 1 gives an estimable version of it though it is somewhat complicated by requiring first step estimation of several distribution and quantile functions.

Step 1: Estimating $F_{Y_{1t}|Y_{0t-1}, D_t=1}$ and $F_{Y_{0t}|Y_{0t-1}, D_t=1}$:

$F_{Y_{1t}|Y_{0t-1}, D_t=1}$ can be estimated directly using local logit as discussed above. For $F_{Y_{0t}|Y_{0t-1}, D_t=1}$, an estimable version of $F_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y')$ is given in Lemma 1. Once again, I estimate the conditional distribution using local logit and by plugging in estimators of the unconditional distributions in Lemma 1. Here,

$$\hat{F}_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0|y') = \Lambda(\hat{\theta}_0(y_0, y') + \hat{\theta}_1(y_0, y')y')$$

where

$$(\hat{\theta}_0(y_0, y'), \hat{\theta}_1(y_0, y')) = \arg \max_{\theta_0, \theta_1} \sum_{i=1}^n \left(\mathbb{1}\{Y_{it-1} \leq z_0\} \log(\Lambda(\theta_0 + \theta_1 Y_{it-2})) \right. \\ \left. + (1 - \mathbb{1}\{Y_{it-1} \leq z_0\}) \log(1 - \Lambda(\theta_0 + \theta_1 Y_{it-2})) \right) K_h(Y_{it-2} - z')$$

where $z_0 = \hat{F}_{Y_{0t-1}|D_t=1}^{-1} \circ \hat{F}_{Y_{0t}|D_t=1}(y_0)$ and $z' = \hat{F}_{Y_{0t-2}|D_t=1}^{-1} \circ \hat{F}_{Y_{0t-1}|D_t=1}(y')$ and n denotes the number

of individuals in the treated group.

Step 2: Compute the Bounds on the Conditional Distribution of the Treatment Effect

To obtain the bounds on the conditional distribution of the treatment effect, one can plug in the above estimates into the results of Lemma 3

$$\hat{F}_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|y') = \sup_y \max\{\hat{F}_{Y_{1t}|Y_{0t-1}, D_t=1}(y|y') - \hat{F}_{Y_{0t}|Y_{0t-1}, D_t=1}(y - \delta|y'), 0\}$$

and

$$\hat{F}_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|y') = 1 + \inf_y \min\{\hat{F}_{Y_{1t}|Y_{0t-1}, D_t=1}(y|y') - \hat{F}_{Y_{0t}|Y_{0t-1}, D_t=1}(y - \delta|y'), 0\}$$

where the sup and inf are taken over a fine grid of possible values for y .

Step 3: Average over the Conditional Bounds of the DoTT

Then, an estimate of the lower bound of the DoTT is given by

$$\hat{F}_{Y_{1t}-Y_{0t}|D_t=1}^L(\delta) = \frac{1}{n} \sum_{i=1}^n \hat{F}_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^L(\delta|Y_{it-1})$$

Similarly, the upper bound on the DoTT is estimated by

$$\hat{F}_{Y_{1t}-Y_{0t}|D_t=1}^U(\delta) = \frac{1}{n} \sum_{i=1}^n \hat{F}_{Y_{1t}-Y_{0t}|Y_{0t-1}, D_t=1}^U(\delta|Y_{it-1})$$

Step 4: Estimate the Bounds on the QoTT

The upper bound on the QoTT is given by inverting the lower bound on the DoTT, and the lower bound on the QoTT is given by inverting the upper bound on the DoTT. Therefore,

$$\text{QoTT}^U(\tau) = \inf\{\delta : \hat{F}_{Y_{1t}-Y_{0t}|D_t=1}^L(\delta) \geq \tau\}$$

and

$$\text{QoTT}^L(\tau) = \inf\{\delta : \hat{F}_{Y_{1t}-Y_{0t}|D_t=1}^U(\delta) \geq \tau\}$$

Estimating the ATT-CPO

Recall that the ATT-CPO is given by

$$\begin{aligned} \text{ATT-CPO}(y') &= E[Y_{1t} - Y_{0t} | Y_{0t-1} = y', D_t = 1] \\ &= E[Y_{1t} | Y_{0t-1} = y', D_t = 1] \end{aligned} \quad (4.1)$$

$$- E[F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(Y_{0t-1})) \mid Y_{0t-2} = F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y')), D_t = 1] \quad (4.2)$$

I estimate the ATT-CPO using local linear kernel estimators. An estimate of the term in Equation 4.1 is given by $\hat{a}_1(y')$ where

$$(\hat{a}_1(y'), \hat{b}_1(y')) = \arg \min_{a_1, b_1} \sum_{i=1}^n [Y_{it} - a_1 - (Y_{it-1} - y')b_1]^2 K_h(Y_{it-1} - y')$$

This is straightforward to estimate as it is simply weighted least squares. Let $\delta_1(y') = [a_1(y'), b_1(y')]'$, \mathcal{Y}_1 be an $n \times 1$ vector with i th component Y_{it} , \mathcal{X}_1 be an $n \times 2$ matrix with i th row given by $(1, Y_{it-1} - y')$, and $\mathcal{K}_1(y')$ be an $n \times n$ diagonal matrix with i th diagonal element given by $K_h(Y_{it-1} - y')$. Then, a closed form expression for the estimate of $\delta_1(y')$ is

$$\hat{\delta}_1(y') = (\mathcal{X}_1' \mathcal{K}_1(y') \mathcal{X}_1)^{-1} \mathcal{X}_1' \mathcal{K}_1(y') \mathcal{Y}_1$$

Estimating the term in Equation 4.2 is more complicated because it depends on distribution functions and quantile functions. These need to be estimated in a first step. With estimates of the distribution functions in hand, Equation 4.1 can be estimated by the solution $\hat{a}_0(y')$ to

$$\begin{aligned} \arg \min_{a_0, b_0} \sum_{i=1}^n & \left(\hat{F}_{Y_{0t}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(Y_{it-1})) - a_0 - \left(Y_{it-2} - \hat{F}_{Y_{0t-2}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(y')) \right) b_0 \right) \\ & \times K_h \left(Y_{it-2} - \hat{F}_{Y_{0t-2}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(y')) \right) \end{aligned}$$

Likewise, there is a closed form solution to this problem. Let $\delta_0(y') = [a_0(y'), b_0(y')]'$, $\hat{\mathcal{Z}}_0$ be an $n \times 1$ vector with i th component $\hat{F}_{Y_{0t}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(Y_{it-1}))$, $\hat{\mathcal{X}}_0$ be an $n \times 2$ matrix with i th row given by $(1, Y_{it-2} - \hat{F}_{Y_{0t-2}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(y')))$, and $\hat{\mathcal{K}}_0(y')$ be an $n \times n$ diagonal matrix with i th diagonal element given by $K_h(Y_{it-2} - \hat{F}_{Y_{0t-2}|D_t=1}^{-1}(\hat{F}_{Y_{0t-1}|D_t=1}(Y_{it-2})))$. Then, a closed form expression for the estimate of $\delta_0(y')$ is

$$\hat{\delta}_0(y') = (\hat{\mathcal{X}}_0' \hat{\mathcal{K}}_0(y') \hat{\mathcal{X}}_0)^{-1} \hat{\mathcal{X}}_0' \hat{\mathcal{K}}_0(y') \hat{\mathcal{Z}}_0$$

4.1 Inference

Inference for the ATT-CPO is relatively straightforward as it is essentially a local linear estimator of a conditional mean with the only complication that it depends on first step estimation of

$F_{Y_{0t}|D_t=1}$, $F_{Y_{0t-1}|D_t=1}$ and $F_{Y_{0t-2}|D_t=1}$. However, each of these will converge at the rate \sqrt{n} which is faster than \sqrt{nh} , the rate that the ATT-CPO will converge. In practice, I choose the bandwidth $h = 1.06sd(Y_{0t-1})n^{-1/4}$ which undersmooths the ATT-CPO thus implying that asymptotically the bias will be negligible at the cost of increasing the variance of the estimator of the ATT-CPO. But this has the advantage of allowing one to use the empirical bootstrap to compute standard errors and not needing to estimate the bias in practice.

Inference for the QoTT is more complicated and is left for future research. The reason that it is more complicated is that it depends on first step non-parametric estimation, second step taking the sup over a function of the first step estimations, and then averaging which is much more non-standard than the ATT-CPO.

5 Job Displacement during the Great Recession

This section studies the effect of job displacement during the Great Recession on yearly earnings of late prime-age workers. Using standard techniques, job displacement is estimated to decrease workers earnings by 39% on average relative to counterfactual earnings had they not been displaced. The size of this effect is 0-25% larger than existing estimates of the effect of job displacement for all workers during severe recessions. The size of the effect is also consistent with the ideas that (i) the effect of job displacement is larger for relatively older workers and (ii) the effect of job displacement is larger during recessions.

Next, this section considers the distributional impacts of job displacement using the techniques developed in the paper. There are two key findings that would not be available without these methods. First, using the panel data methods developed in the paper provides substantially more identifying power for distributional treatment effects such as the QoTT than is available using existing bounds that do not exploit panel data. The reason the bounds are tighter is that relatively strong positive dependence is observed in non-displaced earnings over time for the displaced group of workers in the period before the Great Recession. These bounds are tight enough to rule out the assumption of perfect positive dependence between displaced and non-displaced potential earnings. This result implies that there is more heterogeneity (and potentially much more heterogeneity) in the effect of job displacement than would be implied by the estimate of the QoTT under the assumption of perfect positive dependence. The bounds also imply that some workers have higher earnings after being displaced than they would have had if they not been displaced; this implies that the assumption of Monotone Treatment Response (Manski 1997) is rejected. Second, workers with higher earnings before the recession experience larger decreases in the level of earnings following displacement, but as a fraction of earnings, the earnings loss is roughly constant across workers with differing earnings prior to the recession.

These results can be compared to existing empirical work on job displacement. Broadly speaking, there are two key findings from the job displacement literature: (i) the effect of job displace-

ment on earnings is large, and (ii) the effect of job displacement is persistent. The current paper considers the effect of job displacement on earnings 2-4 years following displacement which is a somewhat shorter period than most existing work. The empirical literature on job displacement finds that workers suffer large earnings losses upon job displacement. To give some examples, Jacobson, LaLonde, and Sullivan (1993) study the effect of job displacement during a deep recession – the recession in the early 1980s. That paper finds that workers lose 40% of their earnings upon displacement and still have 25% lower earnings six years following displacement. Also, it finds little difference in the path of earnings for older, prime-age, and younger workers. Couch and Placzek (2010) study job displacement in the smaller recession in the early 2000s. They find an initial 32% decrease in earnings following displacement, but earnings are only 13% lower six years after displacement. Using Social Security data that covers the entire U.S., Von Wachter, Song, and Manchester (2009) also study the effect of displacement during the early 1980s and find a 30% reduction in earnings upon displacement and earnings still 20% lower up to twenty years following displacement. Because they have data on the entire country, they can compare the effect of displacement including and excluding observations with 0 earnings in a particular year. Not surprisingly, the effect of job displacement is larger when 0 earnings are included, but the path of earnings is very similar – a large dip followed by some recovery but never complete recovery. Kletzer and Fairlie (2003), using NLSY data, find that displaced workers have 11% lower earnings three years after displacement than they would have had if they had not been displaced. That paper uses the same dataset as in the current paper and finds considerably smaller effects of job displacement; however, it considers the period 1984-1993 where the workers are much younger (they would have ranged from 20-36 over those years) and the economy did not experience a deep recession which likely work together to explain the large differences. Stevens (1997), using PSID data, finds that workers initially lose 25% of their earnings following job displacement and have 9% lower earnings ten years later. Using the Displaced Worker Survey, Farber (1997) finds that displaced workers lose 12% of weekly earnings on average following displacement. The effect of job displacement on earnings is larger when there are weak labor market conditions relative to strong labor market conditions (Farber 1997; Davis and Von Wachter 2011) which is relevant for workers displaced during the Great Recession.

There are two potentially important sources of bias in most of the work on job displacement. First, most work limits the sample to individuals who have positive earnings in each period. This may be important because difficulty finding new employment is likely to be a consequence of job displacement. In studies that use state-level administrative data such as Jacobson, LaLonde, and Sullivan (1993) and Couch and Placzek (2010), this choice is made because they are unable to tell whether 0 earnings represents unemployment, leaving the labor force, or moving to another state. Under the condition that more productive displaced workers are more likely to return to work than less productive workers, dropping individuals with no earnings is likely to cause the estimated effect of job displacement to be biased towards 0.

A second well known potential problem is that employers may selectively lay off their least productive workers during recessions (Gibbons and Katz 1991). If this is the case, then comparing these workers to workers that are not displaced may tend to overestimate the effect of job displacement if these workers earnings would not have increased as much as non-displaced workers in the absence of job displacement.

The effect of job displacement may be particularly severe for workers displaced during the Great Recession because of the particularly weak labor market conditions in the period immediately following the recession. From the official beginning of the recession in December 2007 to October 2009, four months after the official end of the recession, the unemployment rate doubled from 5.0% to 10.0% (U.S. Bureau of Labor Statistics 2015b). And during the same period, the economy shed almost 8.4 million jobs (U.S. Bureau of Labor Statistics 2015a). For late prime-age workers, ages 45 to 54, the unemployment rate doubled from 3.6% to 7.1% (U.S. Bureau of Labor Statistics 2015c).

There is recent work on the effect of job displacement during the Great Recession using the Displaced Workers Survey Farber (2015). For all workers, the incidence of job loss was at its highest during the Great Recession compared to all other periods covered by the DWS (1981-present). Roughly, one in six workers report having lost a job. Compared to previous time periods, the rate of reemployment is very low with more workers being reemployed in part time jobs. Interestingly, Farber (2015) finds that, for workers who find full time jobs following job displacement, the effect on earnings during the Great Recession has not been unusually large compared to other periods – about 12%. This provides some evidence that the main channel for a differential effect of job displacement during the Great Recession relative to other periods comes from either failure to find a new job or moving from a full time job to a part time job.

5.1 Data

The data comes from the 1979 National Longitudinal Survey of Youth (NLSY). The main outcome variable is log of yearly earnings in 2011. In 2011, NLSY respondents are between 47 and 54 years old. I limit the sample to individuals who worked at least 1000 hours in 2007 and classify workers as being displaced if they left a job in 2008 or 2009 and the reason given is (i) layoff, job eliminated or (ii) company, office, or workplace closed. This excludes other reasons for leaving a job such as being fired, quitting, moving, or the end of temporary employment.

A main issue in the job displacement literature is how to treat individuals with 0 earnings. Most research drops individuals with 0 earnings (for example, research that uses state-level administrative data cannot tell the difference between actually having 0 earnings and moving to another state). Even most research on job displacement using the PSID or NLSY tends to drop individuals with 0 earnings, then take the log of earnings as the outcome variable in some regression (Stevens 1997; Kletzer and Fairlie 2003). I follow this same approach and restrict the sample to individuals who have at least \$1,000 in earnings in 2003, 2007, and 2011. This sample includes

2,939 individuals, of whom 133 are displaced in 2008 or 2009.

Summary Statistics are provided in Table 1. Displaced workers and non-displaced workers have similar education levels. There are no statistically significant differences in the fraction that have less than high school education, graduated from high school, or graduated from college. There are also negligible differences in gender and in their percentile score on the Armed Forces Qualification Test (AFQT). On the other hand, there are larger differences in race; displaced workers are 12 percentage points more likely to be black (p-value: 0.01).

Table 1 also contains the path of average earnings for displaced and non-displaced workers from 2001-2013. In 2001, average yearly earnings for displaced workers is about \$41,000 while average earnings for non-displaced workers is about \$45,000 (the p-value on the difference being statistically significant is 0.28). From 2001-2007, which are all pre-displacement years, the gap remains roughly constant and does not display any obvious trend over time. However, by 2009, displaced workers have experienced a large decline in average earnings. The gap in 2009 is roughly \$16,000 and almost all of this is explained by the decline in earnings of the group of displaced workers. By 2011, the gap in earnings has increased to just over \$20,000 with almost all of the increase due to increases in average earnings for the group of non-displaced workers.⁹ The gap is still \$22,000 in 2013.

5.2 Baseline Results

In this section, I estimate the average effect of job displacement on the earnings of late prime-age workers using the Change in Changes model. The results indicate that late prime-age workers lose 39% of their earnings on average due to job displacement. This effect is the same or somewhat larger in magnitude compared to estimates of the effect on prime age workers during the deep recession in the early 1980s (Jacobson, LaLonde, and Sullivan 1993; Von Wachter, Song, and Manchester 2009) which are the largest in the literature. This estimate is more than three times as large as the estimated effect of job displacement on all workers during the Great Recession (Farber 2015).¹⁰

I also tried several alternative approaches including Difference in Differences, Difference in Differences where the trend can depend on covariates as in Heckman, Ichimura, and Todd (1997) and Abadie (2005), and the Change in Changes model with covariates. The results are very similar in all cases with estimates ranging from 36% lower earnings to 38% lower earnings.

⁹Recall that workers are displaced in either 2008 or 2009, so earnings in 2009 may mix pre-displacement earnings for workers who are displaced at some point in 2009 with their post-displacement earnings which may account for the smaller earnings gap in 2009 than in 2011 or 2013.

¹⁰The actual difference is probably not as large because the results in Farber (2015) come from workers who worked full time before and after job displacement. This is likely to be important empirically as only 50% of workers are reemployed at the time of their interview and 20% more are employed part time. The estimates in the current paper would not include those that are not reemployed, but it would include those that are employed part time who, by construction, will tend to have lower earnings. Moreover, estimates from the DWS have tended to produce lower estimated effects of job displacement on earnings than estimates from other sources.

5.3 The Distributional Effects of Job Displacement

This section uses the techniques developed earlier in the paper to understand the distributional effects of job displacement. I focus on estimating the QoTT and the ATT-CPO. The QoTT is useful for understanding heterogeneity in the effect of job displacement. It is partially identified in the current application. The ATT-CPO is point identified and is useful for determining whether job displacement has had a relatively larger impact on workers who previously had high or low earnings.

There are three main findings using the methods developed in the current paper. First, the bounds developed in the current paper provide substantially more identifying power than bounds relying on only knowledge of the marginal distributions of potential outcomes. Those bounds are essentially uninformative in the current application. Second, the tighter bounds in the current paper rule out: (i) the assumption of perfect positive dependence across treated and untreated potential outcomes and (ii) the assumption of Monotone Treatment Response (Manski 1997). Finally, estimates of the ATT-CPO (See Figure 6) imply that workers that had higher earnings in the previous period experience larger earnings losses due to job displacement than workers with lower earnings in the previous period. But, as a fraction of earnings, the effect of job displacement is very similar for workers across all initial earnings levels.

Recall that the three key requirements to estimate these distributional treatment effect parameters are (i) access to panel data, (ii) identification of the counterfactual distribution of potential outcomes, and (iii) the Copula Stability Assumption. Thus, as a first step, I need to estimate the counterfactual distribution of potential outcomes – this is what I do next.

Step 1: Estimate the counterfactual distribution of untreated potential outcomes for the treated group

The first task to be accomplished is to estimate the counterfactual distribution of untreated potential outcomes for the treated group – in other words, the unobserved distribution of earnings for the group of displaced workers if they had not been displaced. Knowledge of this distribution, in combination with the distribution of treated potential outcomes for the treated group (which is observed), identifies the QTT. I use the Change in Changes method of Athey and Imbens (2006) though there are a variety of other methods that could be used for the first step estimation including the Panel DID method of Callaway and Li (2017), a version of the Change in Changes model that allows for covariates, or by assuming selection on observables holds after conditioning on a lag of earnings (Firpo 2007). Figure 8 plots the QTT under each of these alternative assumptions. The QTT results are not sensitive to changes in the underlying model for the counterfactual distribution of untreated potential outcomes. This result also implies that the results for parameters that depend on the joint distribution of potential outcomes are not sensitive to the choice of identifying assumption in the first stage.

Estimates of both marginal distributions are presented in Figure 2, and an estimate of the

QTT is presented in Figure 3. The QTT is negative everywhere and statistically significant for all quantiles except the very largest (primarily due to standard errors increasing). The QTT is also increasing across quantiles. It is tempting to interpret Figure 3 as indicating that the effect of job displacement is largest for individuals at the lower part of the income distribution. However, as mentioned above, job displacement appears to cause many individuals to change their rank in the distribution of earnings which makes this interpretation of the QTT invalid. In fact, the QTT is also consistent with the effect of job displacement being very heterogeneous across individuals; that is, upon job displacement some individuals lose a large fraction of their earnings while others experience less severe effects. We find more evidence for the latter interpretation of the QTT in the next section.

Step 2: Estimate parameters that depend on the joint distribution of treated and untreated potential outcomes

Next, I use the techniques presented earlier in the paper to estimate some parameters that depend on the joint distribution of treated and untreated potential outcomes. First, I consider the QoTT. Figure 4 plots (i) bounds on the QoTT under no assumptions on the dependence between potential outcome distributions and (ii) the bounds developed in the current paper.

There are several things to notice from the figure. First, when no information besides the identified marginal distributions is used, the bounds on the QoTT are very wide. For example, the median of the treatment effect is bounded to be between an earnings losses of 65% and an earnings gain of 56%. The effect of displacement among those most affected by displacement, for example the 5th percentile of the QoTT, is estimated to be between a 64% and 98% loss of earnings. The effect of displacement on those least affected by displacement, for example the 95th percentile of the QoTT is bounded between a 18% loss of earnings and a 630% increase in earnings. From these bounds, one is not able to determine much. These bounds indicate that at least 22% of displaced workers have lower earnings than they would have had if they had not been displaced.

Next, under the Copula Stability Assumption, the bounds are indeed tighter. Earnings losses at the 5th percentile are between 86% and 97% which implies that some individuals lose almost all of their earnings due to displacement. The estimates of the QoTT also imply that at least 43% of individuals are worse off from being displaced. Interestingly, one can also conclude that at least 7.0% of individuals have higher earnings after being displaced than they would have had they not been displaced. This type of conclusion was not available without exploiting the Copula Stability Assumption and would imply that the assumption of Monotone Treatment Response (Manski 1997) is not valid in the current case.

Next, Figure 5 plots the QoTT under several assumptions that would lead to point identification. First, it plots the QoTT under perfect positive dependence between Y_{1t} and Y_{0t} . I have argued that this is an especially strong assumption in this case. For example, it essentially restricts any previously high earnings individuals from moving into much lower paying positions

following displacement. This identifying assumption implies the least amount of heterogeneity in the effect of being displaced. At the 5th percentile, individuals lose 82% from being displaced. At the 95th percentile, they lose 17%. At the median, they lose 27%, and this effect is largely constant across most of the interior quantiles. Of course, the no-assumptions bounds cannot rule out perfect positive dependence between Y_{1t} and Y_{0t} , but under the Copula Stability Assumption, perfect positive dependence is ruled out because the bounds imply more heterogeneity than occurs under perfect positive dependence.

Finally, I also plot the results for the case with perfect positive dependence between Y_{0t} and Y_{0t-1} . This assumption results in considerably more heterogeneity in the effect of job displacement than the assumption of perfect positive dependence between Y_{1t} and Y_{0t} . For example, at the 5th percentile, the estimated effect of job displacement is a loss of 92% of earnings. At the median, the estimated effect is 27% lower earnings per year. And at the 95th percentile, earnings are estimated to be 44% higher than they would have been without job displacement. Further, 78% of individuals are estimated to be worse off from job displacement and 22% are estimated to have higher earnings than they would have had they not been displaced. The reason that the bounds in the current paper are close to the estimates of the QoTT under this point identifying assumption is that strong positive dependence, though not perfect positive dependence, is observed between Y_{0t-1} and Y_{0t-2} for the group of displaced workers (Spearman's Rho = 0.78); the dependence between Y_{1t} and Y_{0t-1} is positive but smaller (Spearman's Rho = 0.55).

Average Treatment Effect Conditional on Previous Outcome

Figure 6 provides estimates of the ATT-CPO. The average effect of job displacement conditional on pre-displacement earnings is not statistically significant from the ATT at any value of previous earnings though the estimates are fairly imprecise. The point estimates range from 26% lower earnings on average for workers who had yearly earnings of around \$15,000 in 2007 to 57% lower earnings on average for workers who had yearly earnings of around \$110,000 in 2007.

These results suggest that the effect of job displacement was roughly the same across pre-recession earnings levels as a fraction of earnings lost or perhaps somewhat larger for individuals that initially had higher earnings. In terms of the effect of job displacement on the level of earnings, the effect appears to be larger for individuals with higher initial earnings. The results from the ATT-CPO are quite different from the interpretation of the QTT in Figure 3 where one might be tempted to infer that the effect of job displacement was larger for individuals in the lower part of the earnings distribution.

Does the Copula Stability Assumption hold for job displacement?

Figure 7 plots Spearman's Rho Y_{0s} and Y_{0s-1} in periods before the Great Recession separately for the treated group and untreated group. Here periods are earnings separated by 4 years starting from 1983 and going to 2011. Spearman's Rho is very similar for both groups. Spearman's Rho is increasing from 1987 (around 0.5 for both groups) to 1995 (around 0.75 for both groups) and

is constant from 1995 - 2011. In 1987, individuals in the NLSY were between 23 and 30 years old. In 1995, they were between 31 and 38 years old. Figure 7 suggests that individuals moved around in the distribution of earnings more in their 20s, that there is less income mobility starting in their 30s and that the amount of mobility is the same through their 40s and early 50s. That Spearman's Rho is constant from 1995-2011 provides a strong piece of evidence in favor of the Copula Stability Assumption.

6 Conclusion

This paper has developed techniques to study distributional treatment effect parameters that depend on the joint distribution of potential outcomes. The results depend on three key ingredients: (i) access to at least three periods of panel data, (ii) identification of the marginal distribution untreated potential outcomes for the treated group and (iii) the Copula Stability Assumption which says that the dependence between untreated potential outcomes over time does not change over time. The last of these is the key idea that allows the researcher to exploit having access to panel data to learn about the joint distribution of potential outcomes. This type of idea may also be useful in other cases where the researcher has access to panel data.

Using these methods, I have studied the distributional effects of job displacement during the Great Recession for late prime-age workers. Using standard techniques, I find that these workers lose on average 39% of their yearly earnings following job displacement. Using the techniques developed in the current paper, I find that this average effect masks substantial heterogeneity: some workers lose a very large fraction of their earnings following job displacement though at least some workers have higher earnings following displacement than they would have had they not been displaced. Finally, I also find that workers with initially higher earnings experience larger earnings losses from job displacement than workers with initially lower earnings, but as a fraction of earnings, the average earnings loss is very similar across initial income levels.

References

- [1] Abadie, Alberto. “Semiparametric Difference-in-Differences Estimators”. *The Review of Economic Studies* 72.1 (2005), pp. 1–19.
- [2] Abbring, Jaap and James Heckman. “Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation”. *Handbook of Econometrics* 6 (2007), pp. 5145–5303.
- [3] Athey, Susan and Guido Imbens. “Identification and Inference in Nonlinear Difference-in-Differences Models”. *Econometrica* 74.2 (2006), pp. 431–497.
- [4] Bitler, Marianne P, Jonah B Gelbach, and Hilary W Hoynes. “Welfare Reform and Children’s Living Arrangements”. *Journal of Human Resources* 41.1 (2006), pp. 1–27.
- [5] Callaway, Brantly and Tong Li. “Quantile Treatment Effects in Difference in Differences Models with Panel Data”. Working Paper. 2017.
- [6] Carneiro, Pedro, Karsten Hansen, and James Heckman. “Removing the Veil of Ignorance in assessing the distributional impacts of social policies”. *Swedish Economic Policy Review* 8 (2001), pp. 273–301.
- [7] Carneiro, Pedro, Karsten Hansen, and James Heckman. “Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice”. *International Economic Review* 44.2 (2003), pp. 361–422.
- [8] Carroll, Robert, David Joulfaian, and Mark Rider. “Income Mobility: The Recent American Experience”. Working Paper, Andrew Young School of Policy Studies Research Paper Series. 2007.
- [9] Chernozhukov, Victor, Ivan Fernandez-Val, and Alfred Galichon. “Quantile and probability curves without crossing”. *Econometrica* 78.3 (2010), pp. 1093–1125.
- [10] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. “Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States”. *The Quarterly Journal of Economics* 129.4 (2014), pp. 1553–1623.
- [11] Couch, Kenneth A and Dana W Placzek. “Earnings losses of displaced workers revisited”. *The American Economic Review* (2010), pp. 572–589.
- [12] Davis, Steven J and Till Von Wachter. “Recessions and the Costs of Job Loss”. *Brookings Papers on Economic Activity* 43.2 (Fall) (2011), pp. 1–72.
- [13] Djebbari, Habiba and Jeffrey Smith. “Heterogeneous Impacts in PROGRESA”. *Journal of Econometrics* 145.1 (2008), pp. 64–80.
- [14] Doksum, Kjell. “Empirical probability plots and statistical inference for nonlinear models in the two-sample case”. *The Annals of Statistics* (1974), pp. 267–277.
- [15] Dong, Yingying and Shu Shen. “Testing for Rank Invariance or Similarity in Program Evaluation: The Effect of Training on Earnings Revisited”. Working Paper. 2015.
- [16] Fan, Jianqing, Mark Farnen, and Irene Gijbels. “Local maximum likelihood estimation and inference”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 60.3 (1998), pp. 591–608.
- [17] Fan, Jianqing and Irene Gijbels. *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability* 66. Vol. 66. CRC Press, 1996.

- [18] Fan, Yanqin, Emmanuel Guerre, and Dongming Zhu. “Partial identification of functionals of the joint distribution of potential outcomes”. *Journal of Econometrics* 197.1 (2017), pp. 42–59.
- [19] Fan, Yanqin and Sang Soo Park. “Partial identification of the distribution of treatment effects and its confidence sets”. *Advances in Econometrics* 25 (2009), pp. 3–70.
- [20] Fan, Yanqin and Sang Soo Park. “Sharp bounds on the distribution of treatment effects and their statistical inference”. *Econometric Theory* 26.03 (2010), pp. 931–951.
- [21] Fan, Yanqin and Sang Soo Park. “Confidence intervals for the quantile of treatment effects in randomized experiments”. *Journal of Econometrics* 167.2 (2012), pp. 330–344.
- [22] Farber, Henry S. “The Changing Face of Job Loss in the United States, 1981-1995”. *Brookings Papers on Economic Activity. Microeconomics* (1997), pp. 55–142.
- [23] Farber, Henry S. “Job Loss in the Great Recession and its Aftermath: U.S. Evidence from the Displaced Workers Survey”. Working Paper. 2015.
- [24] Firpo, Sergio. “Efficient Semiparametric Estimation of Quantile Treatment Effects”. *Econometrica* 75.1 (2007), pp. 259–276.
- [25] Firpo, Sergio and Geert Ridder. “Bounds on functionals of the distribution of treatment effects” (2008).
- [26] Frandsen, Brigham R and Lars J Lefgren. “Testing rank similarity”. *Review of Economics and Statistics* (2015).
- [27] Frandsen, Brigham R and Lars J Lefgren. “Partial Identification of the Distribution of Treatment Effects”. Working Paper. 2016.
- [28] Frank, Maurice, Roger Nelsen, and Berthold Schweizer. “Best-possible bounds for the distribution of a sum-a problem of Kolmogorov”. *Probability Theory and Related Fields* 74.2 (1987), pp. 199–211.
- [29] Fréchet, Maurice. “Sur Les Tableaux de Correlation Dont Les Marges Sont Donnees”. *Annals University Lyon: Series A* 14 (1951), pp. 53–77.
- [30] Frölich, Markus. “Non-parametric regression for binary dependent variables”. *The Econometrics Journal* 9.3 (2006), pp. 511–540.
- [31] Gibbons, Robert and Lawrence F Katz. “Layoffs and Lemons”. *Journal of Labor Economics* 9 (1991), pp. 351–380.
- [32] Gottschalk, Peter. “Inequality, Income Growth, and Mobility: The Basic Facts”. *The Journal of Economic Perspectives* (1997), pp. 21–40.
- [33] Hall, Peter, Rodney CL Wolff, and Qiwei Yao. “Methods for estimating a conditional distribution function”. *Journal of the American Statistical Association* 94.445 (1999), pp. 154–163.
- [34] Hansen, Bruce E. “Nonparametric estimation of smooth conditional distributions”. Working Paper: Department of Economics, University of Wisconsin. 2004.
- [35] Heckman, James, Hidehiko Ichimura, and Petra Todd. “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme”. *The Review of Economic Studies* 64.4 (1997), pp. 605–654.

- [36] Heckman, James and Jeffrey Smith. “Evaluating the Welfare State James J. Heckman and Jeffrey Smith”. *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*. 31. Cambridge University Press. 1998, pp. 241–318.
- [37] Heckman, James, Jeffrey Smith, and Nancy Clements. “Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts”. *The Review of Economic Studies* 64.4 (1997), pp. 487–535.
- [38] Hoeffding, Wassily. “Massstabinvariante korrelationstheorie”. *Schriften des Mathematischen Instituts und des Institutes Fur Angewandte Mathematik der Universitat Berlin* 5 (1940), pp. 179–233.
- [39] Hungerford, Thomas L. “US Income Mobility in the Seventies and Eighties”. *Review of Income and Wealth* 39.4 (1993), pp. 403–417.
- [40] Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan. “Earnings losses of displaced workers”. *The American Economic Review* (1993), pp. 685–709.
- [41] Joe, Harry. *Multivariate Models and Multivariate Dependence Concepts*. CRC Press, 1997.
- [42] Joe, Harry. *Dependence Modeling with Copulas*. CRC Press, Boca Raton, FL, 2015.
- [43] Kim, Ju Hyun. “Identifying the Distribution of Treatment Effects under Support Restrictions”. Working Paper. 2014.
- [44] Kletzer, Lori G and Robert W Fairlie. “The long-term costs of job displacement for young adult workers”. *ILR Review* 56.4 (2003), pp. 682–698.
- [45] Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. “Earnings inequality and mobility in the United States: Evidence from social security data since 1937”. *The Quarterly Journal of Economics* 125.1 (2010), pp. 91–128.
- [46] Lehmann, Erich. *Nonparametrics: Statistical Methods Based on Ranks*. 1974.
- [47] Makarov, GD. “Estimates for the Distribution Function of a Sum of Two Random Variables When the Marginal Distributions are Fixed”. *Theory of Probability & its Applications* 26.4 (1982), pp. 803–806.
- [48] Manski, Charles. “Monotone Treatment Response”. *Econometrica* (1997), pp. 1311–1334.
- [49] Nelsen, Roger. *An Introduction to Copulas*. 2nd ed. Springer, 2007.
- [50] Rüschendorf, Ludger. “Random variables with maximum sums”. *Advances in Applied Probability* (1982), pp. 623–632.
- [51] Sen, Amartya. *On Economic Inequality*. Clarendon Press, 1997.
- [52] Sklar, Abe. *Fonctions de répartition à n dimensions et leurs marges*. Publications de L Institut de Statistique de L Universite de Paris, 1959.
- [53] Stevens, Ann Huff. “Persistent Effects of Job Displacement: The Importance of Multiple Job Losses”. *Journal of Labor Economics* (1997), pp. 165–188.
- [54] U.S. Bureau of Labor Statistics. *All Employees: Total Nonfarm Payrolls [PAYEMS]*. 2015.
- [55] U.S. Bureau of Labor Statistics. *Civilian Unemployment Rate [UNRATE]*. 2015.
- [56] U.S. Bureau of Labor Statistics. *Unemployment Rate: 45 to 54 Years [LNU04000093]*. Retrieved from FRED, Federal Reserve Bank of St. Louis on August 9, 2017. 2015.

- [57] Von Wachter, Till, Jae Song, and Joyce Manchester. “Long-term earnings losses due to mass layoffs during the 1982 recession: An analysis using US administrative data from 1974 to 2004”. *unpublished; New York: Columbia University* (2009).
- [58] Williamson, Robert and Tom Downs. “Probabilistic Arithmetic. I. Numerical Methods for Calculating Convolutions and Dependency Bounds”. *International Journal of Approximate Reasoning* 4 (1990), pp. 89–158.

A Proofs

A.1 Proof of Result 1

The joint distribution can be expressed as

$$\begin{aligned}
F_{Y_{1t}, Y_{0t}|D_t=1}(y_1, y_0) &= P(Y_{1t} \leq y_1, Y_{0t} \leq y_0 | D_t = 1) \\
&= P(Y_{1t} \leq y_1, F_{Y_{0t}|D_t=1}(Y_{0t}) \leq F_{Y_{0t}|D_t=1}(y_0) | D_t = 1) \\
&= P(Y_{1t} \leq y_1, F_{Y_{0t-1}|D_t=1}(Y_{0t-1}) \leq F_{Y_{0t}|D_t=1}(y_0) | D_t = 1) \\
&= P(Y_{1t} \leq y_1, Y_{0t-1} \leq F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0)) | D_t = 1)
\end{aligned}$$

which implies the result and where the third equality uses the assumption of perfect positive dependence.

A.2 Proof of Lemma 1

The first part holds under the Copula Stability Assumption as follows

$$\begin{aligned}
F_{Y_{0t}, Y_{0t-1}|D_t=1}(y_0, y') &= C_{Y_{0t}, Y_{0t-1}|D_t=1}(F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-1}|D_t=1}(y')) \\
&= C_{Y_{0t-1}, Y_{0t-2}|D_t=1}\left(F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-1}|D_t=1}(y')\right) \\
&= F_{Y_{0t-1}, Y_{0t-2}|D_t=1}\left(F_{Y_{0t-1}|D_t=1}^{-1} \circ F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-2}|D_t=1}^{-1} \circ F_{Y_{0t-1}|D_t=1}(y')\right)
\end{aligned}$$

where the first equality holds from Sklar's Theorem, the second from the Copula Stability Assumption and the third holds from the definition of a copula.

For the second part, the result follows using essentially the same argument as in the proof of Theorem 4 given below and is therefore omitted.

A.3 Proof of Lemma 2

Lemma 2 follows by an application of the Fréchet-Hoeffding bounds to a conditional bivariate distribution.

A.4 Proof of Lemma 3

Lemma 3 applies the sharp bounds on the difference between random variables with known marginal distributions but unknown copula of Williamson and Downs (1990) (See also: Makarov (1982), Rüschendorf (1982), Frank, Nelsen, and Schweizer (1987), and Fan and Park (2010)) to the difference conditional on the previous outcome.

A.5 Proofs of Theorem 1 and Theorem 2

Theorem 1 and Theorem 2 follow from results in Fan and Park (2010, Section 5) and Fan, Guerre, and Zhu (2017) which derive bounds on the unconditional distribution of the treatment

effect when conditional marginal distributions are known. In those cases, the marginal distributions are conditional on observed covariates X ; in the current paper, the marginal distributions are conditional on a lag of the outcome Y_{0t-1} .

A.6 Proof of Theorem 3

Theorem 3 holds because inverting sharp bounds on a distribution implies sharp bounds on the quantiles (Williamson and Downs 1990; Fan and Park 2010).

A.7 Proof of Proposition 1

To simplify notation throughout the proof, let $F_{1|0} = F_{Y_{1t}|Y_{0t-1}, D_t=1}$ and let $F_{t-1} = F_{Y_{0t-1}|D_t=1}$.

Upper Bound:

Let $F_1^U(y_0, y_1)$ denote the upper bound on the joint distribution given in Theorem 1 when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_1$ and likewise let $F_2^U(y_0, y_1)$ denote the upper bound when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_2$ and under the conditions stated in the proposition. The goal is to show that $F_1^U(y_0, y_1) - F_2^U(y_0, y_1) > 0$ for all $(y_0, y_1) \in \mathcal{Y} \times \mathcal{Y}$.

First, it is straightforward to show that

$$\begin{aligned} F_1^U(y_0, y_1) - F_2^U(y_0, y_1) &= \int_{\mathcal{Y}} (F_1(y_0|y') - F_2(y_0|y')) \mathbb{1}\{F_1(y_0|y') < F_{1|0}(y_1|y'), F_2(y_0|y') < F_{1|0}(y_1|y')\} dF_{t-1}(y') \\ &+ \int_{\mathcal{Y}} (F_1(y_0|y') - F_{1|0}(y_1|y')) \mathbb{1}\{F_1(y_0|y') < F_{1|0}(y_1|y'), F_2(y_0|y') > F_{1|0}(y_1|y')\} dF_{t-1}(y') \\ &+ \int_{\mathcal{Y}} (F_{1|0}(y_1|y') - F_2(y_0|y')) \mathbb{1}\{F_1(y_0|y') > F_{1|0}(y_1|y'), F_2(y_0|y') < F_{1|0}(y_1|y')\} dF_{t-1}(y') \end{aligned} \quad (\text{A.1})$$

Next, $F_{1|0} \prec^{SI} F_1 \prec^{SI} F_2$ implies that there exist unique $y_1^* = y_1^*(y_0, y_1)$ and $y_2^* = y_2^*(y_0, y_1)$ such that

$$F_{1|0}(y_1|y_1^*) = F_1(y_0|y_1^*) \quad \text{and} \quad F_{1|0}(y_1|y_2^*) = F_2(y_0|y_2^*)$$

and

$$\text{for } y' > y_1^*, F_{1|0}(y_1|y') > F_1(y_0|y') \quad \text{for } y' < y_1^*, F_{1|0}(y_1|y') < F_1(y_0|y')$$

and

$$\text{for } y' > y_2^*, F_{1|0}(y_1|y') > F_2(y_0|y') \quad \text{for } y' < y_2^*, F_{1|0}(y_1|y') < F_2(y_0|y')$$

Plugging these into Equation A.1 implies that the difference in Equation A.1 is given by

$$\int_{\mathcal{Y}} (F_1(y_0|y') - F_2(y_0|y')) \mathbb{1}\{y' > y_1^*, y' > y_2^*\} dF_{t-1}(y') \quad (\text{A.2})$$

$$+ \int_{\mathcal{Y}} (F_1(y_0|y') - F_{1|0}(y_1|y')) \mathbb{1}\{y_1^* < y' < y_2^*\} dF_{t-1}(y') \quad (\text{A.3})$$

$$+ \int_{\mathcal{Y}} (F_{1|0}(y_1|y') - F_2(y_0|y')) \mathbb{1}\{y_2^* < y' < y_1^*\} dF_{t-1}(y') \quad (\text{A.4})$$

Case 1: $y_1^* < y_2^*$. In this case, Equation A.4 is equal to 0. Moreover, in Equation A.3, since y' is restricted to be less than y_2^* , $F_{1|0}(y_1|y') < F_2(y_0|y')$ over this range which implies

$$\begin{aligned} F_1^U(y_0, y_1) - F_2^U(y_0, y_1) &> \int_{\mathcal{Y}} (F_1(y_0|y') - F_2(y_0|y')) \mathbb{1}\{y' > y_1^*\} \, dF_{t-1}(y') \\ &= F_2(y_0, y_1^*) - F_1(y_0, y_1^*) > 0 \end{aligned}$$

where the last inequality holds because $F_1 \prec^{SI} F_2 \implies F_1 \prec^C F_2$ (Joe 1997, Theorem 2.12) where $F \prec^C G$ is the concordance ordering and indicates that $F(x_1, x_2) \leq G(x_1, x_2)$ for all x_1 and x_2 and where F and G are joint distributions with the same marginals.

Case 2: $y_1^* > y_2^*$. In this case, Equation A.3 is equal to 0. Also, Equation A.4 is always positive because y' is restricted to be greater than y_2^* . Equation A.2 is then given by

$$\int_{\mathcal{Y}} (F_1(y_0|y') - F_2(y_0|y')) \mathbb{1}\{y' > y_2^*\} \, dF_{t-1}(y') = F_2(y_0, y_2^*) - F_1(y_0, y_2^*) > 0$$

Lower Bound:

Let $F_1^L(y_0, y_1)$ denote the upper bound of the joint distribution given in Theorem 1 when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_1$ and likewise let $F_2^L(y_0, y_1)$ denote the upper bound when $F_{Y_{0t}|Y_{0t-1}, D_t=1} = F_2$. For this part, the goal is to show that $F_2^L(y_0, y_1) - F_1^L(y_0, y_1) > 0$ for all $(y_0, y_1) \in \mathcal{Y} \times \mathcal{Y}$. Similar to the case for the upper bound, one can show that

$$\begin{aligned}
F_2^L(y_0, y_1) - F_1^L(y_0, y_1) &= \int_{\mathcal{Y}} (F_2(y_0|y') - F_1(y_0|y')) \mathbb{1}\{F_{1|0}(y_1|y') + F_1(y_0|y') - 1 > 0, F_{1|0}(y_1|y') + F_2(y_0|y') - 1 > 0\} \\
&\quad - \int_{\mathcal{Y}} (F_{1|0}(y_1|y') + F_1(y_0|y') - 1) \mathbb{1}\{F_{1|0}(y_1|y') + F_1(y_0|y') - 1 > 0, F_{1|0}(y_1|y') + F_2(y_0|y') - 1 < 0\} \\
&\quad + \int_{\mathcal{Y}} (F_{1|0}(y_1|y') + F_2(y_0|y') - 1) \mathbb{1}\{F_{1|0}(y_1|y') + F_1(y_0|y') - 1 < 0, F_{1|0}(y_1|y') + F_2(y_0|y') - 1 > 0\}
\end{aligned} \tag{A.5}$$

Because $F_{1|0}$, F_1 , and F_2 are all stochastically increasing, there exist unique $y_1^\dagger = y_1^\dagger(y_0, y_1)$ and $y_2^\dagger = y_2^\dagger(y_0, y_1)$ such that

$$F_{1|0}(y_1|y_1^\dagger) + F_1(y_0|y_1^\dagger) = 1 \quad \text{and} \quad F_{1|0}(y_1|y_2^\dagger) + F_2(y_0|y_2^\dagger) = 1$$

and

$$\text{for } y' > y_1^\dagger, F_{1|0}(y_1|y') + F_1(y_0|y') < 1 \quad \text{for } y' < y_1^\dagger, F_{1|0}(y_1|y') + F_1(y_0|y') > 1$$

and

$$\text{for } y' > y_2^\dagger, F_{1|0}(y_1|y') + F_2(y_0|y') < 1 \quad \text{for } y' < y_2^\dagger, F_{1|0}(y_1|y') + F_2(y_0|y') > 1$$

One can plug these into Equation A.5 to obtain

$$\begin{aligned}
& F_2^L(y_0, y_1) - F_1^L(y_0, y_1) \\
&= \int_{\mathcal{Y}} (F_2(y_0|y') - F_1(y_0|y')) \mathbb{1}\{y' < y_1^\dagger, y' < y_2^\dagger\}
\end{aligned} \tag{A.6}$$

$$-\int_{\mathcal{Y}} (F_{1|0}(y_1|y') + F_1(y_0|y') - 1) \mathbb{1}\{y_2^\dagger < y' < y_1^\dagger\} \quad (\text{A.7})$$

$$+ \int_{\mathcal{Y}} (F_{1|0}(y_1|y') + F_2(y_0|y') - 1) \mathbb{1}\{y_1^\dagger < y' < y_2^\dagger\} \quad (\text{A.8})$$

Case 1: $y_2^\dagger < y_1^\dagger$

In this case, Equation A.8 is equal to 0. For Equation A.7, $F_{1|0}(y_1|y') < 1 - F_2(y_0|y')$ when $y' > y_2^\dagger$ which implies that

$$\begin{aligned} F_2^L(y_0, y_1) - F_1^L(y_0, y_1) &> \int_{\mathcal{Y}} (F_2(y_0|y') - F_1(y_0|y')) \mathbb{1}\{y' < y_1^\dagger\} \\ &= F_2(y_0, y_1^\dagger) - F_1(y_0, y_1^\dagger) > 0 \end{aligned}$$

which holds because $F_1 \prec^C F_2$ as implied by the assumptions in the proposition and which implies the result.

Case 2: $y_2^\dagger > y_1^\dagger$

In this case, Equation A.7 is equal to 0. Equation A.8 is positive because $F_{1|0}(y_1|y') + F_2(y_0|y') > 1$ for $y' < y_2^\dagger$. And Equation A.6 is equal to $F_2(y_0, y_1^\dagger) - F_1(y_0, y_1^\dagger)$ which is also positive.

A.8 Proof of Theorem 4

To show the result in Theorem 4, it must be shown that

$$\mathbb{E}[Y_{0t}|Y_{0t-1} = y', D_t = 1] = \mathbb{E}[F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(Y_{0t-1})) \mid Y_{0t-2} = F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))]$$

Start with

$$\begin{aligned} \mathbb{E}[Y_{0t}|Y_{0t-1} = y', D_t = 1] &= \int_{\mathcal{Y}} y_0 f_{Y_{0t}|Y_{0t-1}, D_t=1}(y_0 \mid y') \, dy_0 \\ &= \int_{\mathcal{Y}} y_0 \frac{f_{Y_{0t}, Y_{0t-1}, D_t=1}(y_0, y')}{f_{Y_{0t-1}|D_t=1}(y')} \, dy_0 \\ &= \int_{\mathcal{Y}} y_0 c_{Y_{0t}, Y_{0t-1}|D_t=1}(F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-1}|D_t=1}(y')) f_{Y_{0t}|D_t=1}(y_0) \, dy_0 \\ &= \int_{\mathcal{Y}} y_0 c_{Y_{0t-1}, Y_{0t-2}|D_t=1}(F_{Y_{0t}|D_t=1}(y_0), F_{Y_{0t-1}|D_t=1}(y')) f_{Y_{0t}|D_t=1}(y_0) \, dy_0 \\ &= \int_{\mathcal{Y}} y_0 f_{Y_{0t-1}, Y_{0t-2}|D_t=1}(F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0)), F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))) \\ &\quad \times \frac{f_{Y_{0t}|D_t=1}(y_0)}{f_{Y_{0t-1}|D_t=1}(F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0))) \times f_{Y_{0t-2}|D_t=1}(F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y')))} \, dy_0 \\ &= \int_{\mathcal{Y}} y_0 f_{Y_{0t-1}|Y_{0t-2}, D_t=1}(F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0)) | F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))) \\ &\quad \times \frac{f_{Y_{0t}|D_t=1}(y_0)}{f_{Y_{0t-1}|D_t=1}(F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0)))} \, dy_0 \end{aligned}$$

Next, make the substitution $u = F_{Y_{0t-1}|D_t=1}^{-1}(F_{Y_{0t}|D_t=1}(y_0))$ which implies

$$y_0 = F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(u))$$

and

$$dy_0 = \frac{f_{Y_{0t-1}|D_t=1}(u)}{f_{Y_{0t}|D_t=1}(F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(u)))}$$

Plugging these back in implies

$$\begin{aligned} E[Y_{0t}|Y_{0t-1} = y', D_t = 1] \\ &= \int_{\mathcal{Y}} F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(u)) f_{Y_{0t-1}|Y_{0t-2}, D_t=1}(u | F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))) \, du \\ &= E[F_{Y_{0t}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(Y_{0t-1})) \mid Y_{0t-2} = F_{Y_{0t-2}|D_t=1}^{-1}(F_{Y_{0t-1}|D_t=1}(y'))] \end{aligned}$$

A.9 Proof of Proposition 2

Using the same arguments as in Athey and Imbens (2006), one can show that

$$F_{Y_{0t}}(y) = F_{Y_{0t-1}} \circ h_{t-1} \circ h_t^{-1}(y)$$

and

$$F_{Y_{0t-1}}(y') = F_{Y_{0t-2}} \circ h_{t-2} \circ h_{t-1}^{-1}(y')$$

These two imply,

$$\begin{aligned} C_{Y_{0t}, Y_{0t-1}|D_t=1}(u, v) &= P(F_{Y_{0t}|D_t=1}(Y_{0t}) \leq u, F_{Y_{0t-1}|D_t=1}(Y_{0t-1}) \leq v | D_t = 1) \\ &= P(F_{Y_{0t-1}|D_t=1} \circ h_{t-1} \circ h_t^{-1}(Y_{0t}) \leq u, F_{Y_{0t-2}|D_t=1} \circ h_{t-2} \circ h_{t-1}^{-1}(Y_{0t-1}) \leq v | D_t = 1) \\ &= P(F_{Y_{0t-1}|D_t=1} \circ h_{t-1}(C + V_t) \leq u, F_{Y_{0t-2}|D_t=1} \circ h_{t-2}(C + V_{t-1}) \leq v | D_t = 1) \\ &= P(F_{Y_{0t-1}|D_t=1} \circ h_{t-1}(C + V_{t-1}) \leq u, F_{Y_{0t-2}|D_t=1} \circ h_{t-2}(C + V_{t-2}) \leq v | D_t = 1) \\ &= P(F_{Y_{0t-1}|D_t=1}(Y_{0t-1}) \leq u, F_{Y_{0t-2}|D_t=1}(Y_{0t-2}) \leq v | D_t = 1) \\ &= C_{Y_{0t-1}, Y_{0t-2}|D_t=1}(u, v) \end{aligned}$$

where the fourth equality holds because of the additional condition in the proposition which implies that (V_t, V_{t-1}, C) follows the same distribution as (V_{t-1}, V_{t-2}, C) .

B Tables and Figures

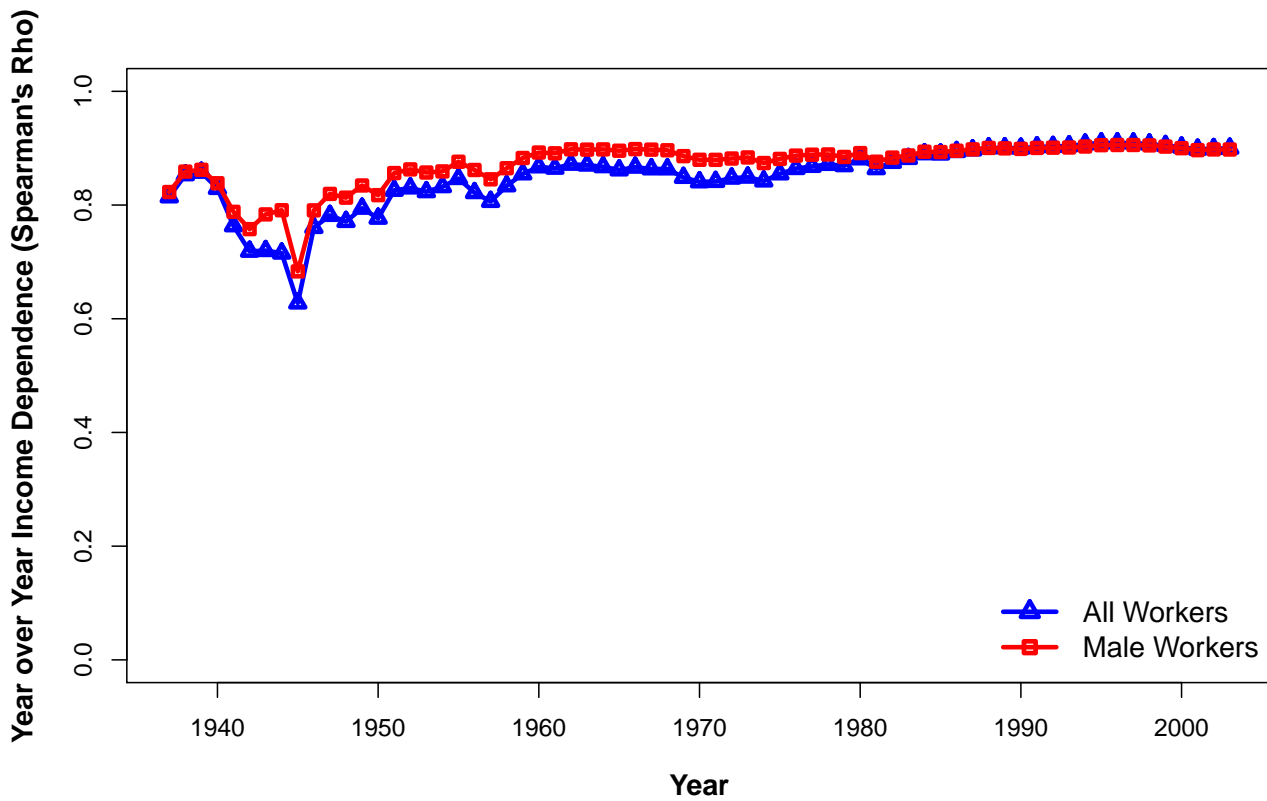
Table 1: Summary Statistics

	Displaced	Non-Displaced	Diff	P-val on Diff
Characteristics				
Less than HS	0.10	0.07	0.03	0.33
High School	0.61	0.61	0.01	0.91
College	0.29	0.32	-0.03	0.50
Hispanic	0.20	0.17	0.03	0.43
Black	0.36	0.24	0.11	0.01
White	0.45	0.59	-0.14	0.00
Female	0.43	0.47	-0.04	0.40
AFQT Score	48.00	50.09	-2.09	0.47
Path of Earnings				
2013 Earnings	43.09	64.92	-21.83	0.00
2011 Earnings	44.70	65.50	-20.80	0.00
2009 Earnings	44.43	60.54	-16.12	0.00
2007 Earnings	55.34	60.39	-5.05	0.35
2005 Earnings	50.07	52.83	-2.77	0.57
2003 Earnings	46.06	49.83	-3.77	0.39
2001 Earnings	40.84	45.22	-4.39	0.28

Notes: Summary statistics for individuals based on whether or not an individual was displaced from their job in 2008 or 2009. The sample used in the table is a restricted version of the main sample as it drops observations that are missing some value of the variables in the table; this restriction results in a sample size of 2,391 (instead of 2939) of which 101 are displaced which amounts to 4.2% of the observations. The variable AFQT Score is the Armed Forces Qualifying Test score normalized to be between 0 and 100. Earnings are in thousands of dollars.

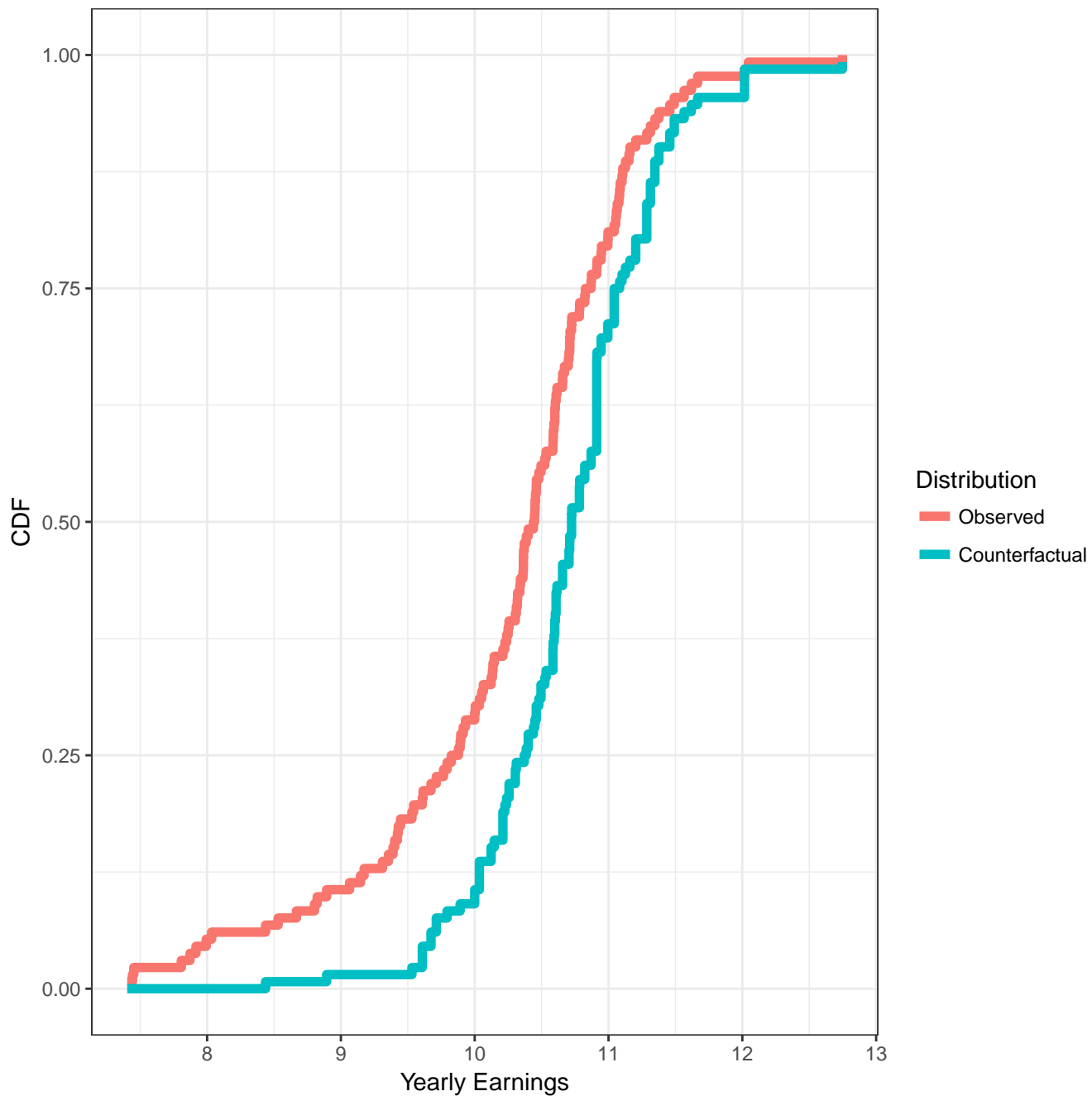
Sources: 1979 National Longitudinal Study of Youth

Figure 1: Rank Correlation (Spearman's Rho) of Year over Year Annual Income Dependence for All Workers and Male Workers from 1937-2003



Notes: The data comes from Kopczuk, Saez, and Song (2010) and replicates part of Figure 4 in that paper.

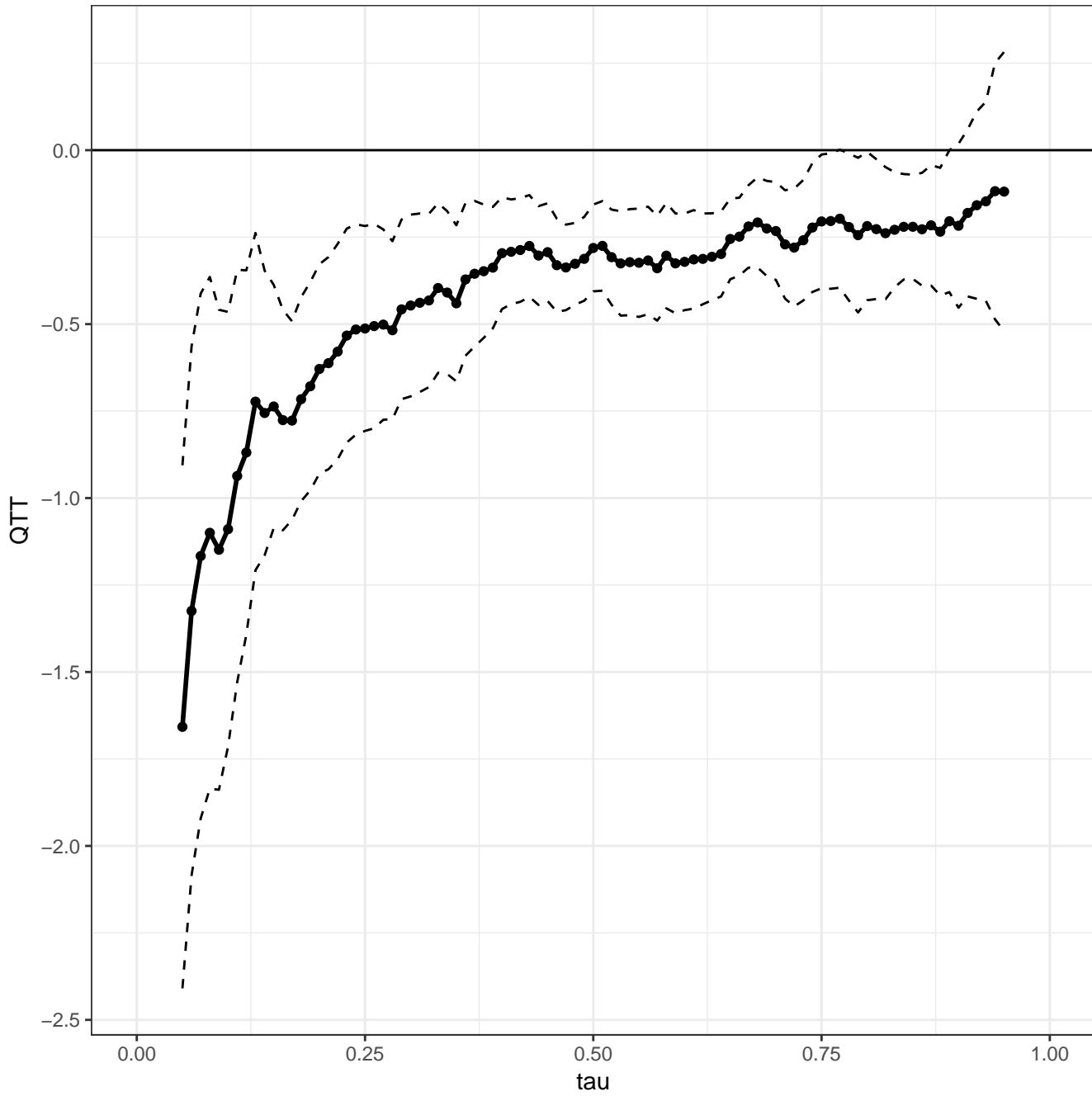
Figure 2: Marginal Distributions of Displaced and Non-displaced Potential Outcomes for the Displaced Group



Notes: This figure provides estimates of the distribution of displaced potential earnings for the treated group and the counterfactual distribution of non-displaced potential earnings for the treated group. The latter is estimated using the Change in Changes model as described in the text.

Sources: 1979 National Longitudinal Survey of Youth

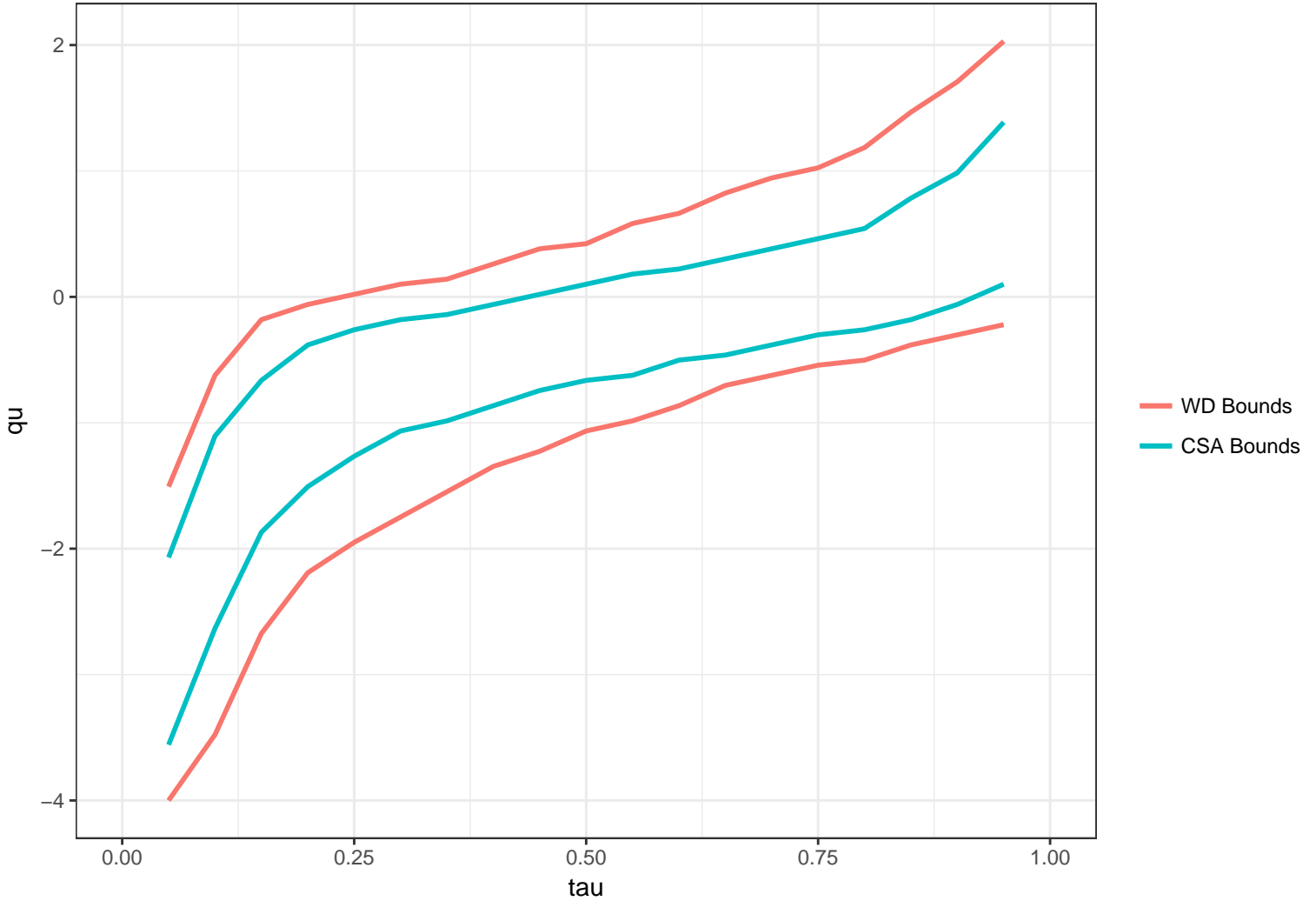
Figure 3: The Quantile Treatment Effect on the Treated



Notes: This figure provides estimates of the QTT of job displacement. The QTT is estimated using the Change in Changes model as described in the text.

Sources: 1979 National Longitudinal Survey of Youth

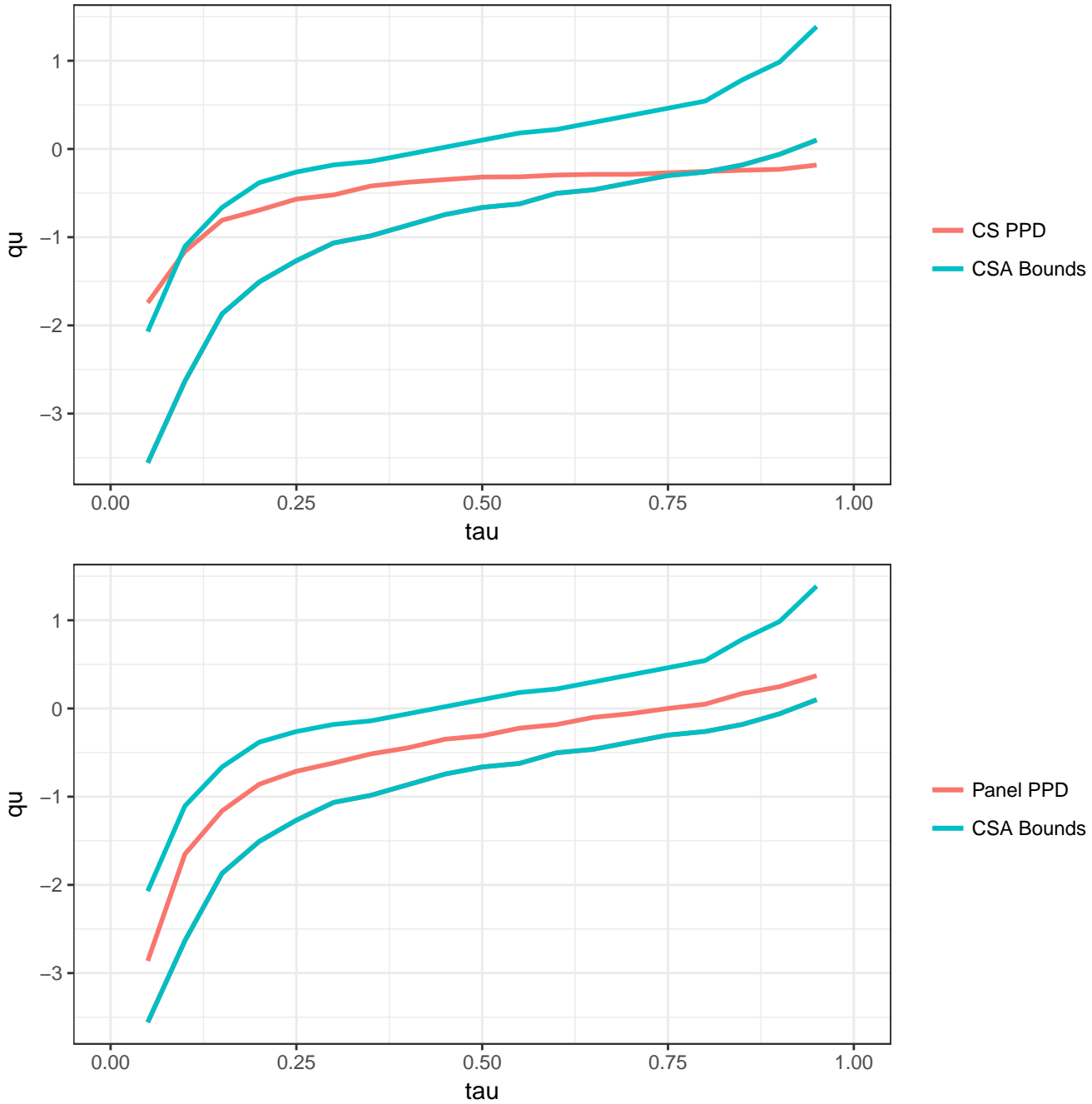
Figure 4: Bounds on the Quantile of the Treatment Effect



Notes: The “WD Bounds” come from using only information about the marginal distributions of displaced and non-displaced potential earnings without applying any restrictions on their dependence. The “CSA Bounds” are the estimates coming from the method in the current paper.

Sources: 1979 National Longitudinal Survey of Youth

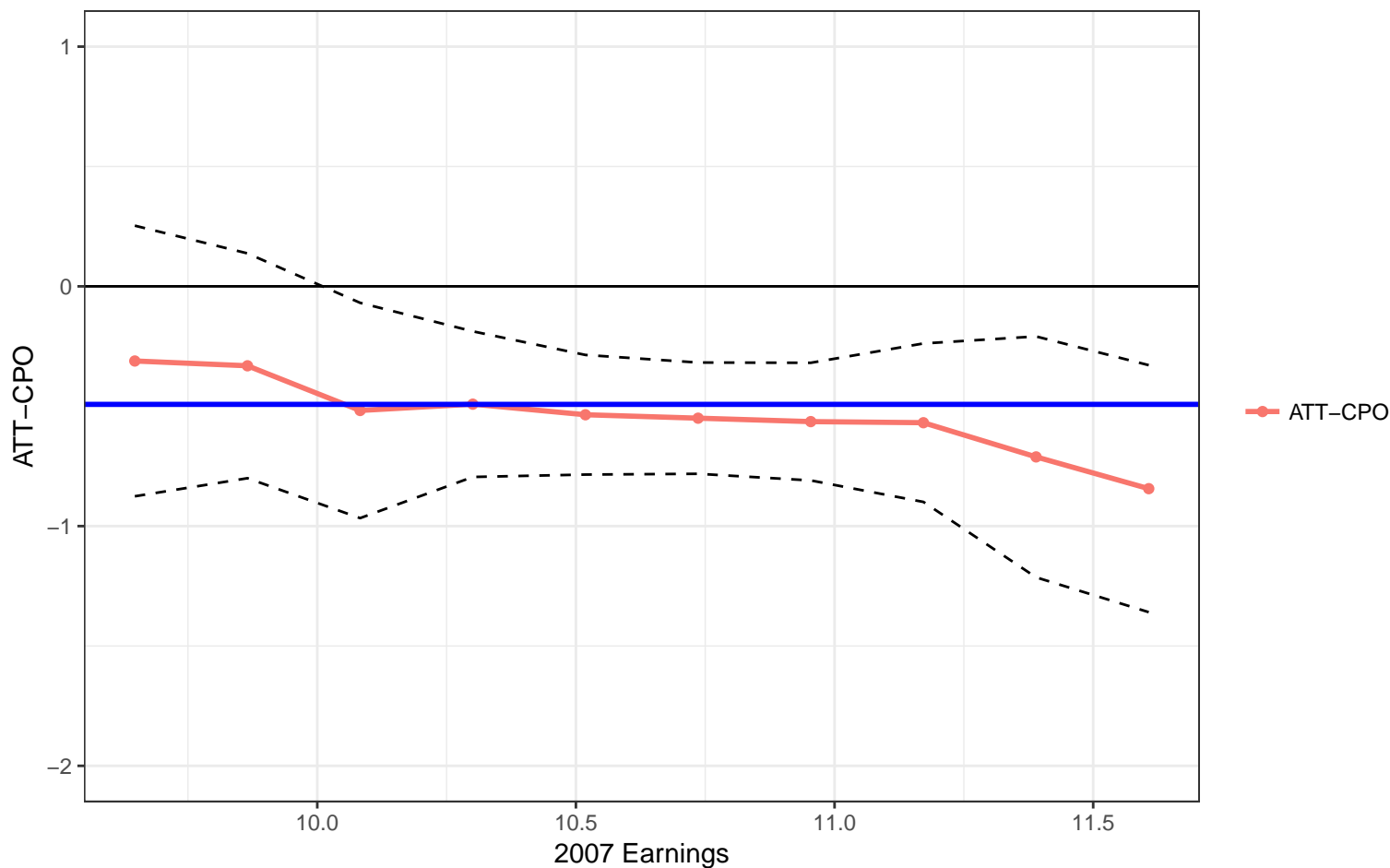
Figure 5: Plots of the QoTT under Rank Invariance Assumptions



Notes: This figure provides plots of the QoTT under the assumption of cross sectional rank invariance (top panel) and over-time rank invariance (bottom panel). The blue lines are the bounds on the QoTT under the Copula Stability Assumption and are the same as in Figure 4.

Sources: 1979 National Longitudinal Survey of Youth

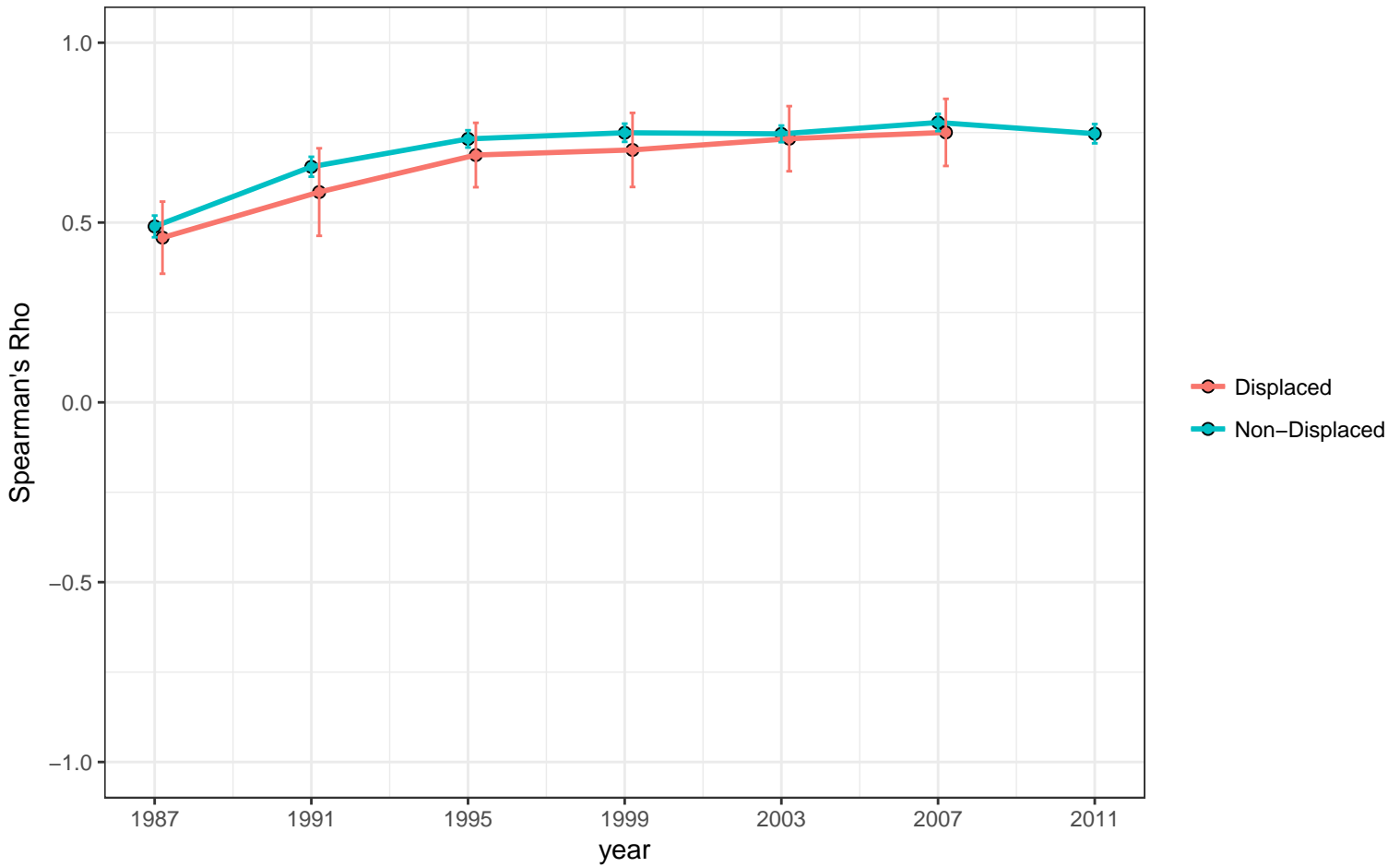
Figure 6: The Average Effect of Job Displacement Conditional on Pre-Recession Earnings (ATT-CPO)



Notes: This figure presents estimates of the average effect of job displacement conditional on pre-recession earnings levels. The blue line is the estimated ATT of job displacement. Pointwise confidence intervals are computed using the bootstrap with 1000 iterations.

Sources: 1979 National Longitudinal Survey of Youth

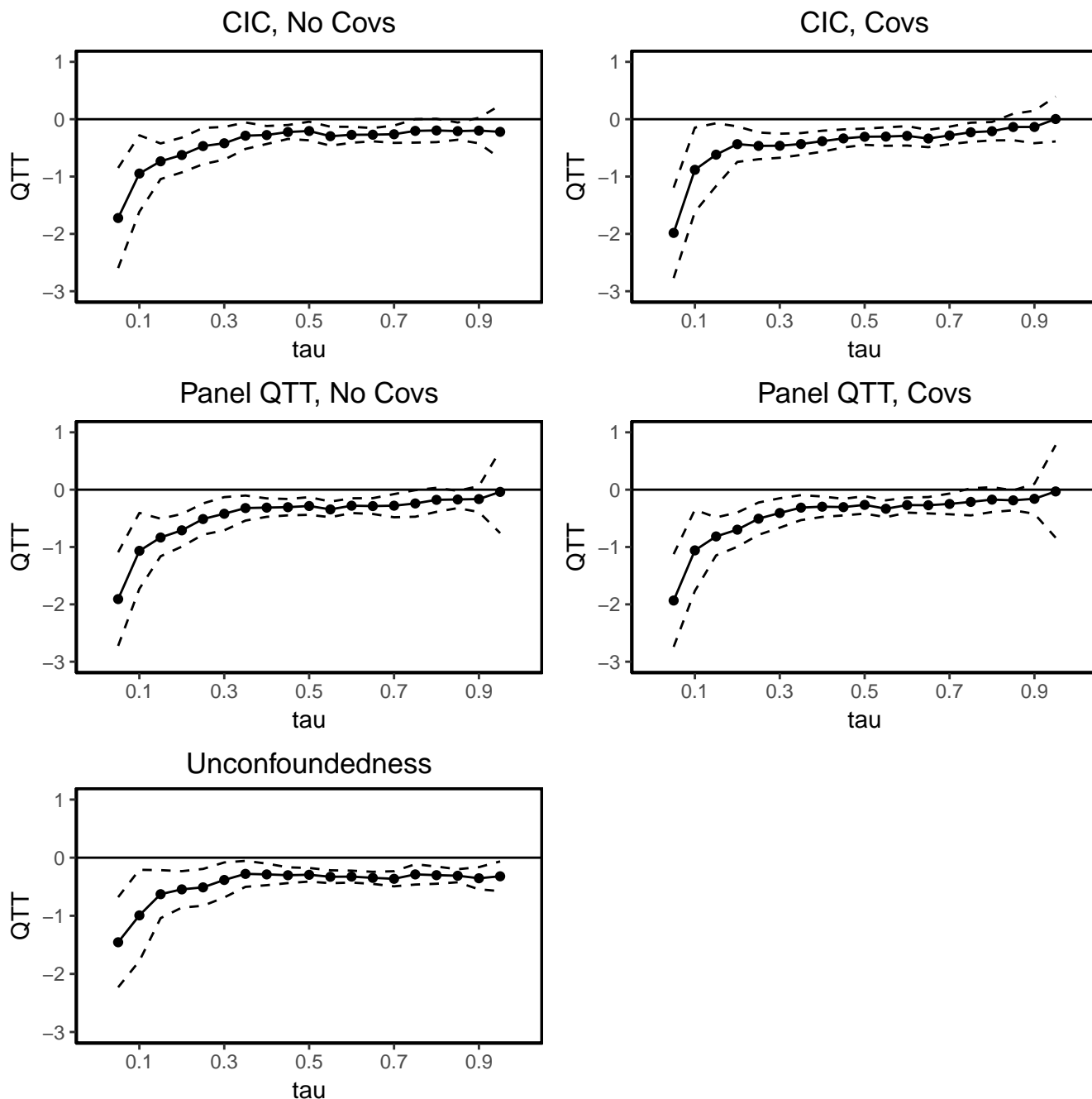
Figure 7: Spearman's Rho for every four year 1983-2007



Notes: This figure provides estimates of Spearman's Rho for the group of displaced workers and the group of non-displaced workers. Spearman's Rho is the correlation of the ranks of earnings in period t and $t - 1$ and depends only on the copula of earnings in period t and period $t - 1$. Standard errors are computed using the block bootstrap with 1000 iterations.

Sources: 1979 National Longitudinal Survey of Youth

Figure 8: Plots of QTTs using Alternative First Step Assumptions



Notes: The figure plots QTTs using alternative assumptions to identify the counterfactual distribution of non-displaced potential earnings for the group of displaced workers. The panel “CIC, No Covs” provides estimates of the QTT using the Change in Changes model with no covariates; these are the same results as presented in Figure 3 up to differences in the underlying samples. The panel “CIC, Covs” includes covariates in the Change in Changes models by first regressing outcomes in each period on the X variables and then running the CIC model on the residuals as suggested in Athey and Imbens (2006). The panel “PQTT, No Covs” uses the Panel QTT method developed in Callaway and Li (2017) without covariates and the “Panel QTT, Covs” uses the same method after adjusting for covariates. The panel “Unconfoundedness” estimates the QTT under the assumption of unconfoundedness using the method developed in Firpo (2007). The last two estimates require a first stage estimation of the propensity score. It is estimated using logit and includes dummy variables for less than high school, high school, or college education; Hispanic, black, or white race; gender; and AFQT score. The unconfoundedness results also include the log of earnings in 2007 as an additional control. The sample includes 2,667 observations for which all of the variables used in each specification are available. This is a subset of the data used in the main results.

Sources: 1979 National Longitudinal Survey of Youth