

Supplementary Appendix: Bounds on Distributional Treatment Effect Parameters using Panel Data with an Application on Job Displacement

Brantly Callaway*

August 27, 2019

This Supplementary Appendix contains a number of additional results for “Bounds on Distributional Treatment Effect Parameters using Panel Data with an Application on Job Displacement.” The first section contains some Monte Carlo simulations to assess the finite sample properties of the estimators discussed in the paper. The second section contains a number of additional theoretical results as well as additional regularity conditions for particular estimators used in the paper. The third section contains some additional results for the application in the paper on job displacement.

SA Monte Carlo Simulations

For the first set of results, I consider the finite sample performance of estimators of the upper bound of the QoTT. To keep things simple, I consider the case where the distribution of $(Y_{0t}, Y_{0t-1})|D = 1$ is known rather than needing to be estimated in a first step. In particular, I consider the case where $(Y_{1t}, Y_{0t}|D = 1) \sim N(0, V_1)$ and $(Y_{0t}, Y_{0t-1}|D = 1) \sim N(0, V_0)$ with

$$V_j = \begin{pmatrix} 1 & \rho_j \\ \rho_j & 1 \end{pmatrix}$$

for $j = 1, 2$.

With a (bivariate) Gaussian distribution, the results in Chen and Fan (2006) and Bouyé and Salmon (2009) imply that

$$P(Y_{jt} \leq y | Y_{0t-1} = y') = \Phi \left(\frac{(y - \rho_j y')}{\sqrt{1 - \rho_j^2}} \right) \quad (\text{SA.1})$$

*Department of Economics, University of Mississippi, 225 Holman Hall, University, MS 38677.
Email: bmcallow@olemiss.edu

Table SA.1: Monte Carlo Simulations

	N=100			N=500			N=1000		
	10%	50%	90%	10%	50%	90%	10%	50%	90%
<u>$\rho_0 = 0$</u>									
Actual	0.249	1.373	3.301						
Bias	-0.259	-0.194	-0.334	-0.096	-0.103	-0.189	-0.062	-0.081	-0.165
MAD	0.225	0.193	0.321	0.096	0.096	0.193	0.064	0.096	0.161
RMSE	0.300	0.258	0.407	0.118	0.131	0.218	0.081	0.101	0.185
<u>$\rho_0 = 0.5$</u>									
Actual	0.056	1.309	3.172						
Bias	-0.284	-0.215	-0.350	-0.115	-0.099	-0.183	-0.056	-0.081	-0.164
MAD	0.289	0.225	0.321	0.129	0.096	0.193	0.064	0.096	0.161
RMSE	0.335	0.275	0.422	0.143	0.126	0.212	0.100	0.097	0.179
<u>$\rho_0 = 0.9$</u>									
Actual	-0.779	0.859	2.691						
Bias	-0.168	-0.185	-0.331	-0.063	-0.084	-0.186	-0.068	-0.065	-0.168
MAD	0.193	0.193	0.321	0.064	0.096	0.193	0.064	0.064	0.161
RMSE	0.264	0.241	0.380	0.108	0.109	0.206	0.092	0.081	0.180

Notes: This table presents results from Monte Carlo experiments using the DGP given in Equation (SA.1). Each section in the table provides estimates across different values of ρ_0 which is the correlation between Y_{0t} and Y_{0t-1} . The rows labeled “Actual” give the actual upper bound on the QoTT for across different quantiles. These change with different values of ρ_0 . “MAD” stands for median absolute deviation, and “RMSE” stands for root mean squared error. The columns provide simulated estimates of the bias, MAD, and RMSE at the 10th percentile, median, and 90th percentile of the treatment effect and using 1000 Monte Carlo simulations.

where Φ is the cdf of a standard normal random variable. Then, the bounds on the QoTT in Theorem 3 (in the main text) can be simulated quite accurately.

For each of the below simulations, I consider the case where $\rho_1 = 0$, and I vary $\rho_0 \in \{0, 0.5, 0.9\}$, and I report the bias, median absolute deviation, and root mean squared error for the upper bound on the 10th percentile, median, and 90th percentile of the treatment effect. I vary the number of observations between 100, 500, and 1000 and randomly assign half of them to being in the treated group.

The results of these Monte Carlo simulations are available in Table SA.1. The main takeaways are as follows. First, across all values of ρ_0 , the performance of the estimators of the upper bound of the QoTT improves with larger sample sizes. Second, the estimators tend to be downward biased overall with the bias decreasing with the sample size. Similar results (which are not shown) hold for the lower bound on the QoTT – estimates of the lower bound on the QoTT tend to be upward biased.

This is not surprising because the upper bound on the QoTT comes from inverting the lower bound on the DoTT – the lower bound on the DoTT is likely to tend to be downward biased because of sampling variation and taking the infimum in Lemma 3 (in the main text); see the related discussion in Manski and Pepper (2000) and Chernozhukov, Lee, and Rosen (2013). Third, the performance of the estimator of the QoTT is somewhat better overall for the median than for the 10th- or 90th-percentiles. Finally, there are not systematic differences in the performance of the estimator across different values of ρ_0 even though larger values of ρ_0 lead to substantially tighter bounds.

For the second set of results, I consider how tight the bounds are under different data generating processes. Here, I use the same DGP as given above. Under this DGP, $Y_{1t}|Y_{0t-1}, D = 1 \sim N(\rho_1 Y_{0t-1}, (1 - \rho_1^2))$; similarly, $Y_{0t}|Y_{0t-1}, D = 1 \sim N(\rho_0 Y_{0t-1}, (1 - \rho_0^2))$. Building on the results in Fan and Park (2010) and Fan and Wu (2010), it holds that bounds on the conditional DoTT are given by

$$F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D=1}^L(\delta|y') = \Phi\left(\frac{\sigma_1 s - \sigma_0 t}{\rho_0^2 - \rho_1^2}\right) + \Phi\left(\frac{\sigma_1 t - \sigma_0 s}{\rho_0^2 - \rho_1^2}\right) - 1$$

$$F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D=1}^U(\delta|y') = \Phi\left(\frac{\sigma_1 s + \sigma_0 t}{\rho_0^2 - \rho_1^2}\right) - \Phi\left(\frac{\sigma_1 t + \sigma_0 s}{\rho_0^2 - \rho_1^2}\right) + 1$$

where Φ is the cdf of a random variable that follows a standard normal distribution and where $\sigma_1^2 := (1 - \rho_1^2)$, $\sigma_0^2 := (1 - \rho_0^2)$, $s = \delta - (\rho_1 - \rho_0)y'$, and $t = \sqrt{s^2 + (\sigma_1^2 - \sigma_0^2) \log\left(\frac{\sigma_1^2}{\sigma_0^2}\right)}$; this holds in the case when $\rho_1 \neq \rho_0$. When $\rho_1 = \rho_0$, the bounds are given by

$$F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D=1}^L(\delta|y') = \mathbb{1}\{\delta \geq (\rho_1 - \rho_0)y'\} \left(2\Phi\left(\frac{\delta - (\rho_1 - \rho_0)y'}{2\sigma}\right) - 1\right)$$

$$F_{Y_{1t}-Y_{0t}|Y_{0t-1}, D=1}^U(\delta|y') = \mathbb{1}\{\delta \geq (\rho_1 - \rho_0)y'\} + \mathbb{1}\{\delta < (\rho_1 - \rho_0)y'\} 2\Phi\left(\frac{\delta - (\rho_1 - \rho_0)y'}{2\sigma}\right)$$

and the bounds on the *DoTT* itself (and corresponding bounds on the QoTT) are given by averaging the conditional bounds over the distribution of $Y_{0t-1}|D = 1$.

Bounds on *QoTT*(0.5) and *QoTT*(0.9) are presented in Table SA.2. There are a few main patterns to notice. First, at least in this case, the bounds have roughly equal length for different values of τ . The bounds get tighter as there is stronger dependence between either Y_{0t} and Y_{0t-1} or Y_{1t} and Y_{0t-1} . Perhaps most interestingly, having very strong dependence between one pair of random variables tightens the bounds relatively more than having a moderate amount of dependence between both pairs of random variables. To see this, notice that the bounds are substantially tighter in the case when, for example, $\rho_0 = 0.9$ and $\rho_1 = 0$ than in the case where $\rho_0 = 0.5$ and $\rho_1 = 0.5$.

Table SA.2: Bounds on QoTT

		0.00		0.50		0.90		0.99	
		ρ_1	τ	0.5	0.9	0.5	0.9	0.5	0.9
ρ_0	τ								
	0.5								
	0.9								
	0.5	2.683	3.544	2.583	3.203	1.702	1.922	0.701	0.741
	0.9	2.583	3.223	2.322	3.063	1.582	1.882	0.681	0.741
0.5	1.722	1.902	1.582	1.882	1.161	1.542	0.561	0.681	
0.9	0.701	0.741	0.681	0.761	0.561	0.681	0.380	0.501	

Notes: The table provides bounds on $QoTT(\tau)$ for $\tau \in \{0.5, 0.9\}$ using the DGP given in this section. Each cell in the table provides $QoTT^U(\tau) - QoTT^L(\tau)$ for different values of the correlation parameters – this is the width of the identified set for different values of the parameters. Each row contains different values of ρ_0 – the correlation between Y_{0t} and Y_{0t-1} . Columns contain different values of ρ_1 – the correlation between Y_{1t} and Y_{0t-1} as well as different results across different values of τ . (Results for $\tau = 0.1$ are not reported because, due to the symmetry of the DGP, the results are the same as for $\tau = 0.9$).

SB Supplementary Asymptotic Results

In the first part of this section, I provide some additional low-level assumptions on the preliminary estimators used in the main part of the text. In cases where a researcher used alternative first step estimators, these regularity conditions would need to be adjusted. The second set of results are on using change-in-changes (Athey and Imbens (2006) and Melly and Santangelo (2015)) in the first step to identify and estimate $F_{Y_{0t}|X, D=1}$. As discussed earlier, other approaches could be used to identify this counterfactual distribution, but the second part of this section supplies additional details for the method that I actually use in the main paper.

SB.1 Additional Assumptions for First Step Estimators

The first two assumptions are technical conditions used in deriving the main asymptotic results in the paper.

Assumption SB.1 (Compact Support).

For all $(s, d) \in \{t, t-1, t-2\} \times \{0, 1\}$, \mathcal{Y}_{ds} and \mathcal{X}_d , which denote the supports of Y_s and X conditional on $D = d$, are compact subsets of \mathbb{R} .

Assumption SB.2 (Continuously Distributed Outcomes).

(i) For all $(s, d) \in \{t, t-1, t-2\} \times \{0, 1\}$, Y_s is continuously distributed conditional on X and $D = d$ with conditional density $f_{Y_d|X, D=d}(y|x)$ that is uniformly bounded away from 0 and ∞ and uniformly continuous in $(y, x) \in \mathcal{Y}_{ds}\mathcal{X}_d$.

(ii) Conditional on Y_{0t-1} , X , and $D = 1$, Y_{1t} and Y_{0t} are continuously distributed with conditional densities $f_{Y_{1t}|Y_{0t-1}, X, D=1}(y|y', x)$ and $f_{Y_{0t}|Y_{0t-1}, X, D=1}(y|y', x)$ that are uniformly bounded away from 0

and ∞ and uniformly continuous in (y, y', x) on their supports.

The next assumption provides additional regularity conditions for the proposed distribution regression estimators of $F_{Y_{1t}|Y_{0t-1}, X, D=1}$ and $F_{Y_{0t}|Y_{0t-1}, X, D=1}$ which are also standard in the literature on distribution regression. First, for $W_{t-1} = (Y_{0t-1}, X^\top)^\top$ and $W_{\Gamma_{20}} = (\Gamma_{20}(Y_{0t-2}, X), X^\top)^\top$, define

$$M_1(y) := E \left[\frac{\lambda(W_{t-1}^\top \beta_1(y))^2}{\Lambda(W_{t-1}^\top \beta_1(y))(1 - \Lambda(W_{t-1}^\top \beta_1(y)))} W_{t-1} W_{t-1}^\top \middle| D = 1 \right] \quad (\text{SB.1})$$

and

$$M_0(y) := E \left[\frac{\lambda(W_{\Gamma_{20}}^\top \beta_0(y))^2}{\Lambda(W_{\Gamma_{20}}^\top \beta_0(y))(1 - \Lambda(W_{\Gamma_{20}}^\top \beta_0(y)))} W_{\Gamma_{20}} W_{\Gamma_{20}}^\top \middle| D = 1 \right] \quad (\text{SB.2})$$

Assumption SB.3 (Distribution Regression).

(i) $E[\|W_{t-1}\|^2 | D = 1] < \infty$ and $E[\|W_{\Gamma_{20}}\|^2 | D = 1] < \infty$.

(ii) The minimum eigenvalues of $M_1(y)$ and $M_0(y)$, which are defined in Equations (SB.1) and (SB.2), are uniformly bounded away from zero.

Next, define $p_d := P(D = d)$ and $p_d(x) := P(D = d | X = x)$. The following assumptions are needed for the particular first step estimators that I use in the application.

Assumption SB.4 (Overlap).

$p_1 > 0$ and, for all $x \in \mathcal{X}_1$, $p_1(x) < 1$.

This assumption is standard in the treatment effects literature. The first part says that there are some treated individuals; the second part says that for any possible values of the covariates for the treated group, there is a positive probability that they there do not participate in the treatment. This guarantees that, for individuals in the treated group, one can find ‘‘matches’’ with the same characteristics. Because all the parameters that are considered in the paper are conditional on being in the treated group, I do not require that that the propensity score, $P(D = 1 | X)$, be bounded away from 0.¹

The next assumption is an additional condition for the first step quantile regression estimators and is standard in the literature on quantile regression.

Assumption SB.5 (First Step Quantile Regression).

(i) For $d \in \{0, 1\}$, $E[\|X\|^{2+\varepsilon} | D = d] < \infty$ for some $\varepsilon > 0$.

(ii) For $\{s, d\} \in \{0, 1\} \times \{t, t-1, t-2\}$, the minimum eigenvalues of $J_{s,d}(u)$, which is defined in Equation (B.1), are uniformly bounded away from zero.

¹This can potentially be important in applications like job displacement where job displacement may be much less common for individuals with particular characteristics.

SB.2 Additional Details for Distribution Regression

Next, in this section, I discuss the results of Chernozhukov, Fernandez-Val, and Melly (2013) which apply directly to estimating $F_{Y_{1t}|Y_{0t-1}, X, D=1}$ because this is just a distribution regression of an observed outcome on some observed covariates. Here, I suppose that $F_{Y_{1t}|Y_{0t-1}, X, D=1}(y|y', x) = \Lambda(w^\top \beta_1(y))$ for some known link function Λ with derivative λ and where $w = (y', x^\top)^\top$. Define

$$H(z) = \frac{\lambda(z)}{\Lambda(z)(1 - \Lambda(z))} \quad (\text{SB.3})$$

and let h denote the derivative of H . Let $\Psi^1(\beta)$ and $\hat{\Psi}^1(\beta)$ be defined as in Equations (B.7) and (B.8) but for estimating $F_{Y_{1t}|Y_{0t-1}, X, D=1}$. That is, they are the population and sample first order conditions for estimating β_1 .

Using the same arguments as in Chernozhukov, Fernandez-Val, and Melly (2013), one can show that

$$\sqrt{n}(\hat{\Psi}^1(\beta_1) - \Psi^1(\beta_1)) \rightsquigarrow \mathbb{Z}_1$$

where

$$\Psi^1(\beta) = E\left[\psi_{y;\beta}^1(Y_t, W_{t-1}, D)\right] \quad \text{and} \quad \hat{\Psi}^1(\beta) = \frac{1}{n} \sum_{i=1}^n \psi_{y;\beta}^1(Y_{it}, W_{it-1}, D_i)$$

with

$$\psi_{y;\beta}^d(Y, W, D) = \frac{\mathbb{1}\{D = d\}}{p_d} (\Lambda(W^\top \beta(y)) - \mathbb{1}\{Y \leq y\}) H(W^\top \beta(y)) W \quad (\text{SB.4})$$

where \mathbb{Z}_1 is a tight, mean zero Gaussian process with covariance function

$$V_{\mathbb{Z}_1}(\tilde{y}_1, \tilde{y}_2) := E\left[\psi_{\tilde{y}_1; \beta_1}^1(Y_t, W_{t-1}, D) \psi_{\tilde{y}_2; \beta_1}^1(Y_t, W_{t-1}, D)^\top\right]$$

Continuing to follow the same arguments as in Chernozhukov, Fernandez-Val, and Melly (2013), it further holds that

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \rightsquigarrow -M_1(\cdot)^{-1} \mathbb{Z}_1$$

and

$$\sqrt{n}(\hat{F}_{Y_{1t}|Y_{0t-1}, X, D=1} - F_{Y_{1t}|Y_{0t-1}, X, D=1}) \rightsquigarrow \mathbb{G}_1 := \lambda(w^\top \beta_1(y)) w^\top M_1(y)^{-1} \mathbb{Z}_1 \quad (\text{SB.5})$$

Finally, I estimate $F_{Y_{0t-1}, X|D=1}$ using the empirical cdf, i.e.,

$$\hat{F}_{Y_{0t-1}, X|D=1}(y', x) = \frac{1}{n} \sum_{i=1}^n \psi_{(y', x)}^{ecdf}(Y_{it-1}, X_i, D_i)$$

where

$$\psi_{(y', x)}^{ecdf}(Y, X, D) := \frac{D}{p_1} \mathbb{1}\{Y \leq y', X \leq x\}$$

Noting that $F_{Y_{0t-1}, X|D=1}(y', x) = E[\psi_{(y', x)}^{ecdf}(Y, X, D)]$, it follows immediately that

$$\sqrt{n}(\hat{F}_{Y_{0t-1}, X|D=1} - F_{Y_{0t-1}, X|D=1}) \rightsquigarrow \mathbb{G}_3 \quad (\text{SB.6})$$

where \mathbb{G}_3 is a tight, mean zero Gaussian process with covariance function given by

$$\begin{aligned} & V_3(\tilde{y}'_1, \tilde{x}_1, \tilde{y}'_2, \tilde{x}_2) \\ &= E \left[(\psi_{(\tilde{y}'_1, \tilde{x}_1)}^{ecdf}(Y_{t-1}, X, D) - F_{Y_{0t-1}, X|D=1}(\tilde{y}'_1, \tilde{x}_1)) (\psi_{(\tilde{y}'_2, \tilde{x}_2)}^{ecdf}(Y_{t-1}, X, D) - F_{Y_{0t-1}, X|D=1}(\tilde{y}'_2, \tilde{x}_2))^\top \right] \end{aligned}$$

SB.3 Additional Details for Change-in-Changes

Under Assumption 2 (in the main text), the distribution of untreated potential outcomes for individuals in the treated group is identified. In the application, I used change-in-changes (Athey and Imbens (2006) and Melly and Santangelo (2015)) to identify this distribution. In particular, in this setup,

$$\begin{aligned} F_{Y_{0t}|X, D=1}(y|x) &= \phi(F_{Y_{0t-1}|X, D=1}, F_{Y_{0t-1}|X, D=0}, F_{Y_{0t}|X, D=0})(y, x) \\ &:= F_{Y_{0t-1}|X, D=1}(F_{Y_{0t-1}|X, D=0}^{-1}(F_{Y_{0t}|X, D=0}(y|x)|x)|x) \end{aligned} \quad (\text{SB.7})$$

where all the terms on the right hand side of Equation (SB.7) are identified. It is natural then to estimate the distribution of untreated potential outcomes for the treated group by

$$\begin{aligned} \hat{F}_{Y_{0t}|X, D=1}(y|x) &= \phi(\hat{F}_{Y_{0t-1}|X, D=1}, \hat{F}_{Y_{0t-1}|X, D=0}, \hat{F}_{Y_{0t}|X, D=0})(y, x) \\ &= \hat{F}_{Y_{0t-1}|X, D=1}(\hat{F}_{Y_{0t-1}|X, D=0}^{-1}(\hat{F}_{Y_{0t}|X, D=0}(y|x)|x)|x) \end{aligned}$$

The limiting process for each of the estimated distributions is given in Equation (B.2), so all that remains to be show is that the function ϕ is Hadamard differentiable. Next, I provide the limiting process for $\sqrt{n}(\hat{F}_{Y_{0t}|X, D=1} - F_{Y_{0t}|X, D=1})$. This is closely related to Melly and Santangelo (2015) though my method of proof is somewhat different and my main result expands one of their intermediate results.

Proposition SB.1. *Let $\mathbb{D} = l^\infty(\mathcal{Y}_{1t-1}\mathcal{X}_1) \times l^\infty(\mathcal{Y}_{0t-1}\mathcal{X}_1) \times l^\infty(\mathcal{Y}_{0t}\mathcal{X}_1)$ and consider the map $\phi : \mathbb{D}_\phi \subset$*

$\mathbb{D} \mapsto l^\infty(\bar{\mathcal{Y}}_{0t}\mathcal{X}_1)$ given by

$$\phi(F) := F_1 \circ F_2^{-1} \circ F_3$$

for $F := (F_1, F_2, F_3) \in \mathbb{D}_\phi$ where $\mathbb{D}_\phi := \mathbb{E}^3$ where \mathbb{E} denotes the set of all conditional distribution functions with conditional density function that is uniformly bounded from above and bounded away from zero. Then, the map ϕ is Hadamard differentiable at $F_0 = (F_{10}, F_{20}, F_{30}) \in \mathbb{D}$ with derivative given by

$$\phi'_{F_0}(\lambda) = \lambda_1 \circ F_{20}^{-1} \circ F_{30} + f_{10}(F_{20}^{-1} \circ F_{30}) \frac{\lambda_3 - \lambda_2 \circ F_{20}^{-1} \circ F_{30}}{f_{20}(F_{20}^{-1} \circ F_{30})}$$

tangentially to \mathbb{D}_ϕ in $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{D}_\phi$.

Proof. Let $\mathbb{D}_{\phi_2} = \mathbb{E} \times \mathbb{E}^{-1} \times \mathbb{E}$ where \mathbb{E}^{-1} is the space of inverse functions in \mathbb{E} . Consider the maps $\phi_1 : \mathbb{D}_\phi \mapsto \mathbb{D}_{\phi_2}$ and $\phi_2 : \mathbb{D}_{\phi_2} \mapsto l^\infty(\bar{\mathcal{Y}}_{0t}\mathcal{X}_1)$ given by

$$\phi_1(F) = (F_1, F_2^{-1}, F_3) \quad \text{and} \quad \phi_2(\Gamma) = (\Gamma_1 \circ \Gamma_2 \circ \Gamma_3)$$

for $F = (F_1, F_2, F_3) \in \mathbb{D}_\phi$ and $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3) \in \mathbb{D}_{\phi_2}$. Notice that $\phi(F) = \phi_2(\phi_1(F))$. Next, the map ϕ_1 is Hadamard differentiable at F_0 with derivative given by

$$\phi'_{1,F_0}(\lambda) = \left(\lambda_1, -\frac{\lambda_2}{f_{20}} \circ F_{20}^{-1}, \lambda_3 \right)$$

(see, for example, van der Vaart and Wellner (1996, Lemma 3.9.23(ii))). Next, the map ϕ_2 is Hadamard differentiable at Γ_0 in $\gamma \in \mathbb{D}_{\phi_2}$ with derivative given by

$$\phi'_{2,\Gamma_0}(\gamma) = \gamma_1 \circ \Gamma_{20} \circ \Gamma_{30} + \Gamma'_{1,\Gamma_{20} \circ \Gamma_{30}} \gamma_2 \circ \Gamma_{30} + \Gamma'_{1,\Gamma_{20} \circ \Gamma_{30}} \Gamma'_{2,\Gamma_{30}} \gamma_3$$

which follows using a similar argument as in van der Vaart and Wellner (1996, Lemma 3.9.27). Further, by the chain rule for Hadamard differentiable functions and for $\lambda \in \mathbb{D}_\phi$,

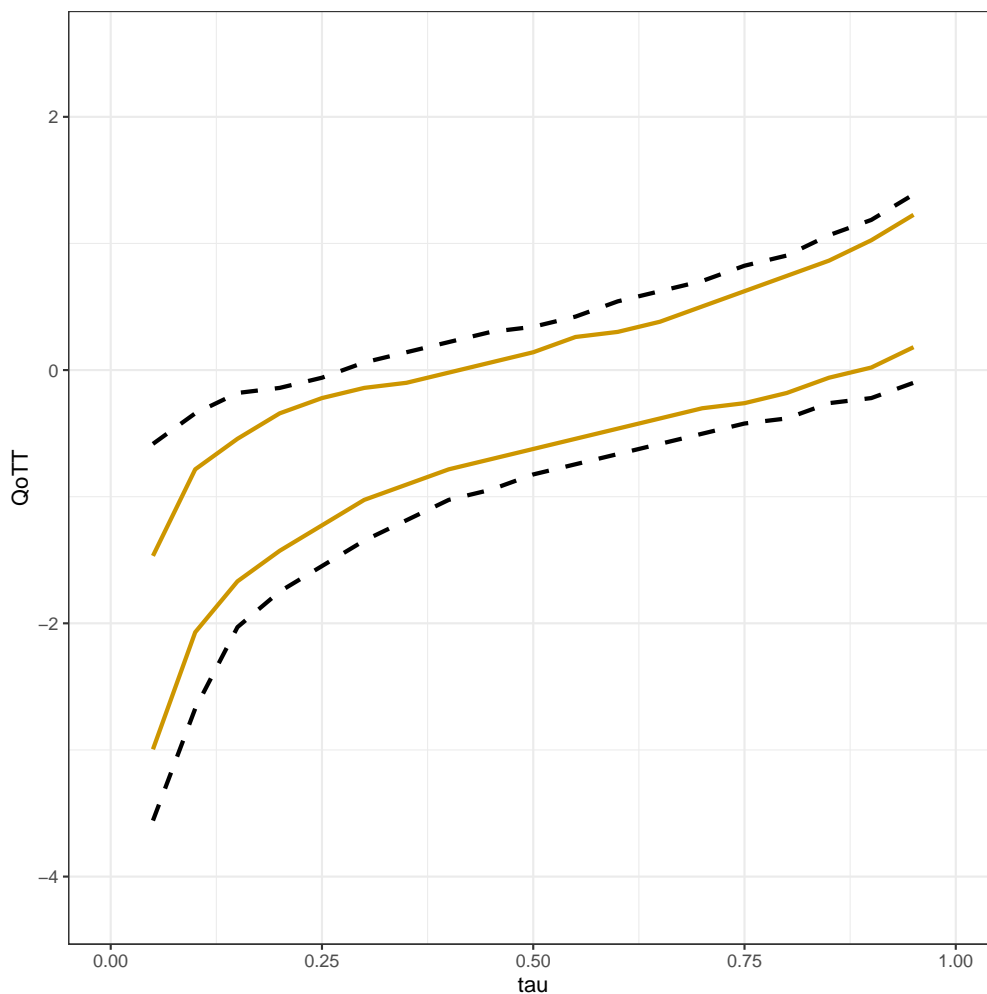
$$\begin{aligned} \phi'_{F_0}(\lambda) &= \phi'_{2,\phi_1(F_0)} \circ \phi'_{1,F_0}(\lambda) \\ &= \phi'_{2,(F_{10}, F_{20}^{-1}, F_{30})} \left(\lambda_1, -\frac{\lambda_2}{f_{20}} \circ F_{20}^{-1}, \lambda_3 \right) \end{aligned}$$

which implies the result by plugging into the expression for $\phi'_{2,\Gamma_0}(\gamma)$ and because the derivative of F_{10} is f_{10} and the derivative of F_{20}^{-1} is $1/(f_{20} \circ F_{20}^{-1})$. □

SC Additional Results from the Application on Job Displacement

This section contains several additional results for the application on job displacement. First, it includes estimates of the bounds on the QoTT when (i) the bounds are tightened using the Copula Stability Assumption and covariates and (ii) the bounds are tightened using covariates but not the Copula Stability Assumption. Figure 6 in the main text provides point estimates of each of these, but Figures SC.1 and SC.2 additionally provide 95% confidence intervals for each of the bounds. Figure SC.3 provides alternative first step estimators of $F_{Y_{0t}|D=1}$. These are largely similar to each other indicating that the bounds on the main parameters of interest in the paper are not sensitive to the choice of identification argument used for this counterfactual distribution.

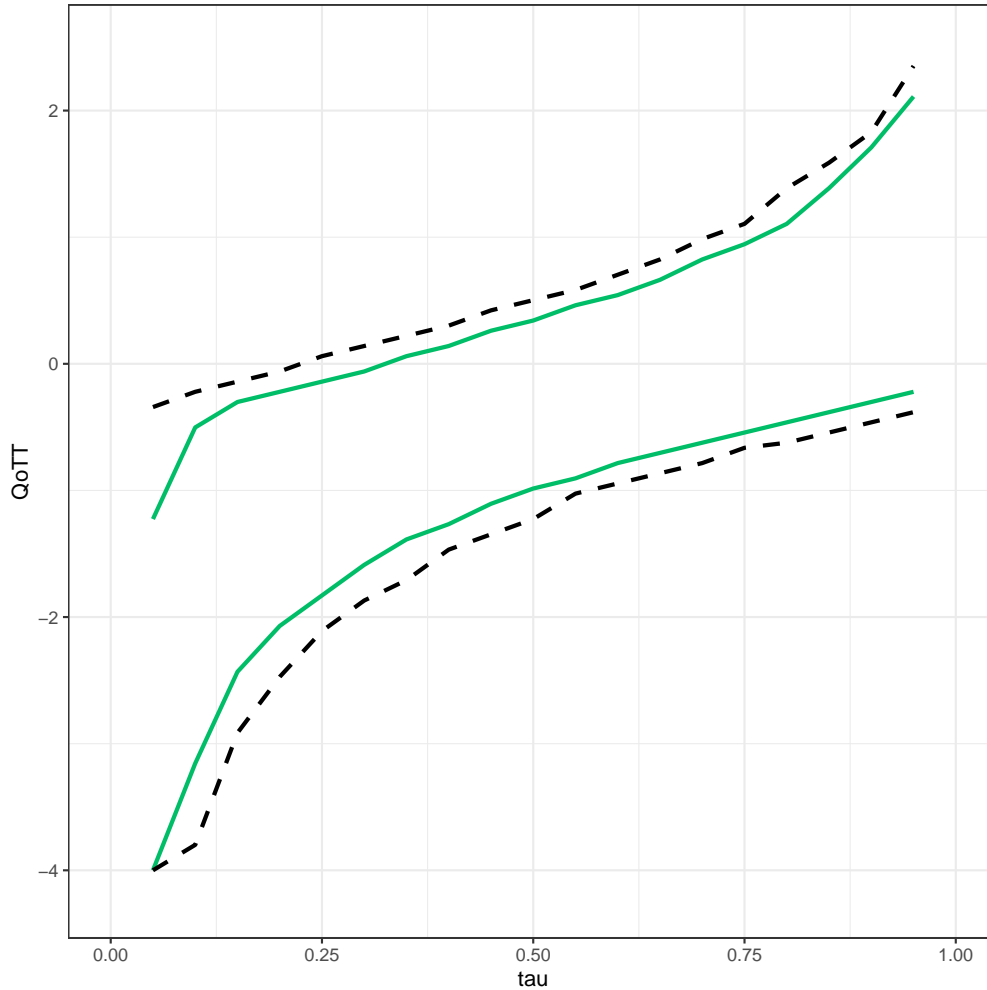
Figure SC.1: Bounds on the Quantile of the Treatment Effect under the Copula Stability Assumption with Covariates



Notes: These are bounds that come from using the method developed in the current paper under the Copula Stability Assumption and through tightening bounds using covariates. The scale of the y-axis is in log points. Most of the reported results in the text convert log points into percentage changes (see Footnote 19). The dotted lines provide 95% confidence intervals for the estimated lower and upper bounds using the numerical bootstrap as discussed in the text.

Sources: 1979 National Longitudinal Survey of Youth

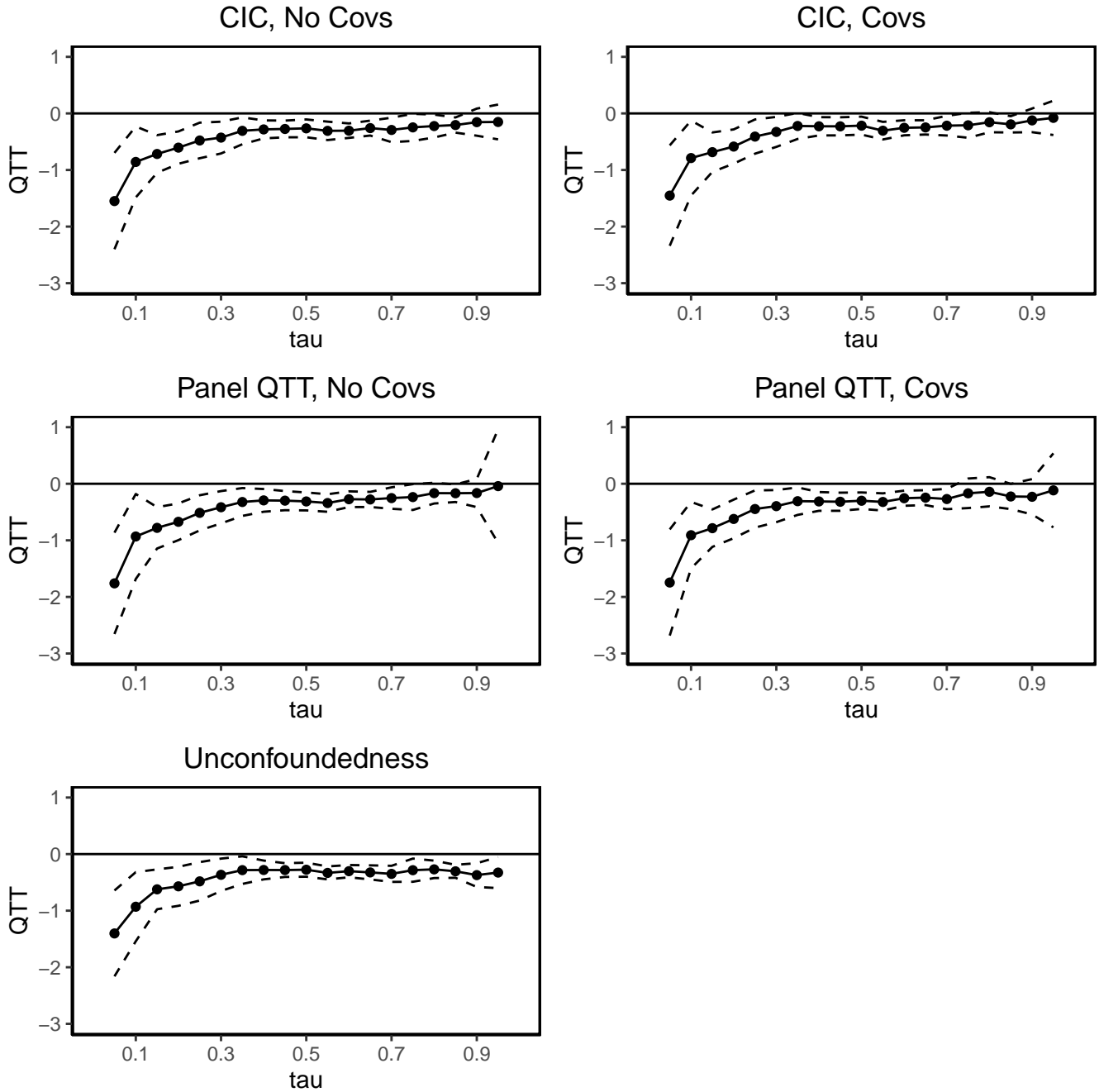
Figure SC.2: Bounds on the Quantile of the Treatment Effect using Covariates



Notes: These are bounds that come from using available covariates to tighten the bounds but do not employ the Copula Stability Assumption. The scale of the y-axis is in log points. Most of the reported results in the text convert log points into percentage changes (see Footnote 19). The dotted lines provide 95% confidence intervals for the estimated lower and upper bounds using the numerical bootstrap as discussed in the text.

Sources: 1979 National Longitudinal Survey of Youth

Figure SC.3: 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. The panel “CIC, Covs” includes covariates in the Change in Changes models using the approach of Melly and Santangelo (2015) that uses first step quantile regression estimators. The panel “Panel QTT, No Covs” uses the Panel QTT method developed in Callaway and Li (2019) 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 a logit model and includes dummy variables for less than high school, high school, or college education; Hispanic, black, or white race; and gender. The unconfoundedness results also include the log of earnings in 2007 as an additional control.

Sources: 1979 National Longitudinal Survey of Youth

References

- [1] Athey, Susan and Guido Imbens. “Identification and inference in nonlinear difference-in-differences models”. *Econometrica* 74.2 (2006), pp. 431–497.
- [2] Bouyé, Eric and Mark Salmon. “Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets”. *The European Journal of Finance* 15.7-8 (2009), pp. 721–750.
- [3] Callaway, Brantly and Tong Li. “Quantile treatment effects in difference in differences models with panel data”. Working Paper. 2019.
- [4] Chen, Xiaohong and Yanqin Fan. “Estimation of copula-based semiparametric time series models”. *Journal of Econometrics* 130.2 (2006), pp. 307–335.
- [5] Chernozhukov, Victor, Ivan Fernandez-Val, and Blaise Melly. “Inference on counterfactual distributions”. *Econometrica* 81.6 (2013), pp. 2205–2268.
- [6] Chernozhukov, Victor, Sokbae Lee, and Adam M Rosen. “Intersection bounds: Estimation and inference”. *Econometrica* 81.2 (2013), pp. 667–737.
- [7] 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.
- [8] Fan, Yanqin and Jisong Wu. “Partial identification of the distribution of treatment effects in switching regime models and its confidence sets”. *The Review of Economic Studies* 77.3 (2010), pp. 1002–1041.
- [9] Firpo, Sergio. “Efficient semiparametric estimation of quantile treatment effects”. *Econometrica* 75.1 (2007), pp. 259–276.
- [10] Manski, Charles and John Pepper. “Monotone instrumental variables: with an application to the returns to schooling”. *Econometrica* 68.4 (2000), pp. 997–1010.
- [11] Melly, Blaise and Giulia Santangelo. “The changes-in-changes model with covariates”. Working Paper. 2015.
- [12] van der Vaart, Aad W and Jon A Wellner. *Weak Convergence and Empirical Processes*. Springer Science & Business Media, 1996.