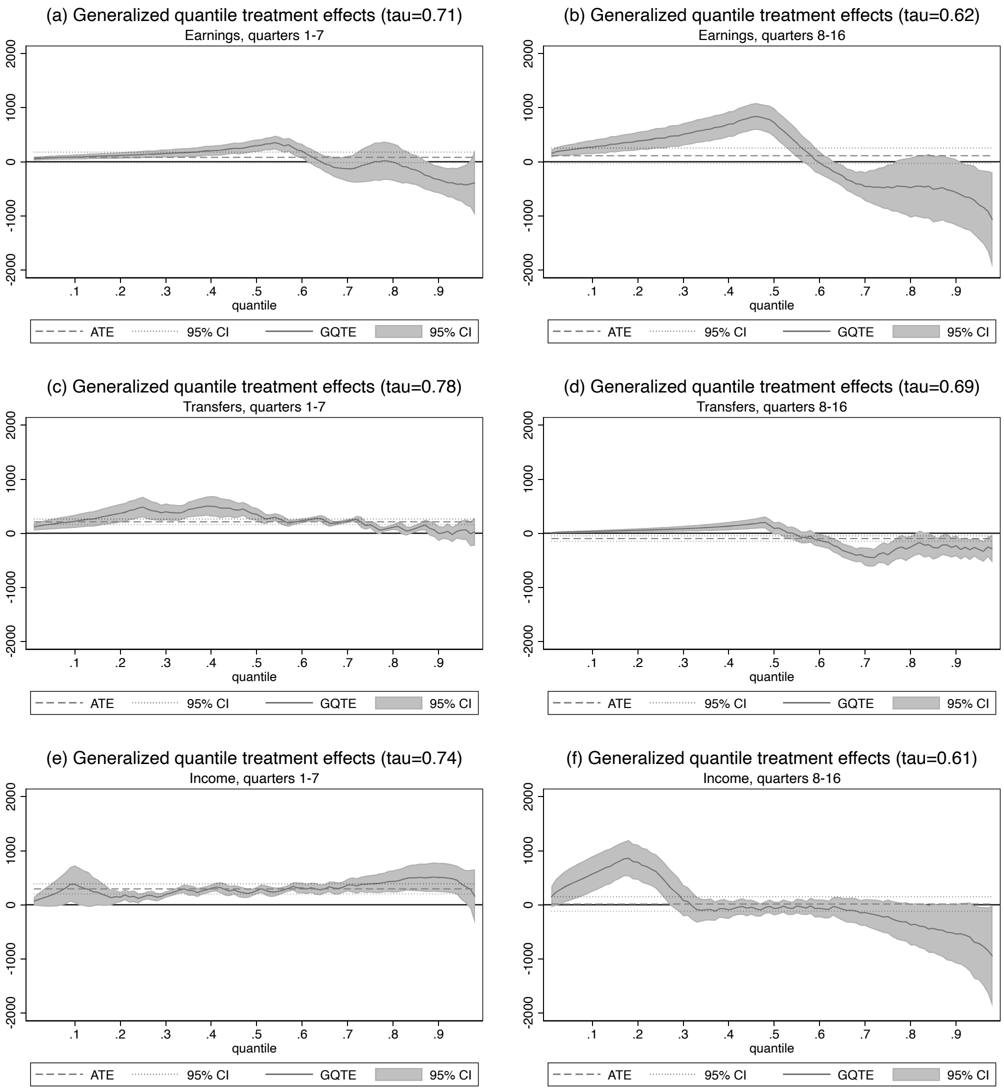
Figure 4: Generalized Quantile Treatment Effects (GQTE)



Weighted results; unrestricted analysis sample. Confidence intervals based on clustered bootstrap standard errors (100 replications) to account for repeated observations. Top percentile not included due to high sampling variability.

Table 3: Winners and Losers: Proportion of Positive and Negative Treatment Effects

	Earnings				Transfers				Income			
	Q1-7		Q8-16		Q1-7		Q8-16		Q1-7		Q8-16	
	(1) QTE	(2) GQTE	(3) QTE	(4) GQTE	(5) QTE	(6) GQTE	(7) QTE	(8) GQTE	(9) QTE	(10) GQTE	(11) QTE	(12) GQTE
$\widehat{\tau}$	1	0.71	1	0.62	1	0.78	1	0.69	1	0.74	1	0.61
% > 0												
Total	34.3	68.7	36.4	59.6	77.8	97.0	1.0	54.5	85.9	99.0	43.4	31.3
$p < 0.05$	31.3	61.6	27.3	56.6	64.6	83.8	0.0	49.5	66.7	81.8	23.2	24.2
$p < 0.01$	22.2	60.6	22.2	55.6	60.6	82.8	0.0	46.5	59.6	75.8	16.2	21.2
$p < 0.001$	18.2	58.6	20.2	53.5	55.6	78.8	0.0	10.1	52.5	71.7	6.1	15.2
% < 0												
Total	12.1	31.3	15.2	40.4	0.0	3.0	47.5	45.5	0.0	1.0	26.3	68.7
$p < 0.05$	1.0	9.1	0.0	22.2	0.0	0.0	33.3	37.4	0.0	0.0	7.1	6.1
$p < 0.01$	0.0	4.0	0.0	7.1	0.0	0.0	26.3	26.3	0.0	0.0	2.0	0.0
$p < 0.001$	0.0	0.0	0.0	3.0	0.0	0.0	17.2	15.2	0.0	0.0	0.0	0.0

Note: Weighted results; unrestricted analysis sample. $\widehat{\tau}$ represents the assumed or estimated rank correlation coefficient between potential treatment and control outcomes. Proportion of positive and negative treatment effects: 'Total' refers to the overall proportion of treatment effects, irrespective of their significance levels. Proportions of treatment effects denoted by $p < 0.05$, $p < 0.01$ and $p < 0.001$ are statistically significant at the 5%, 1% and 0.1% levels, respectively. Bootstrap standard errors (100 replications) were clustered to account for repeated observations.

Table 4: Total Gains and Losses, and Standard Deviation of Treatment Effects

	Earnings				Transfers				Income			
	Q1-7		Q8-16		Q1-7		Q8-16		Q1-7		Q8-16	
	(1) QTE	(2) GQTE	(3) QTE	(4) GQTE	(5) QTE	(6) GQTE	(7) QTE	(8) GQTE	(9) QTE	(10) GQTE	(11) QTE	(12) GQTE
$\hat{\tau}$	1	0.71	1	0.62	1	0.78	1	0.69	1	0.74	1	0.61
Gains (USD '000s)												
Total	276	259	382	649	517	604	2	102	742	704	190	383
$p < 0.05$	269	254	354	640	493	589	0	96	704	621	137	352
$p < 0.01$	203	251	317	634	480	586	0	92	668	580	105	328
$p < 0.001$	184	241	301	616	464	575	0	37	633	560	43	260
Losses (USD '000s)												
Total	53	179	43	516	0	1	246	287	0	24	108	439
$p < 0.05$	7	80	0	344	0	0	215	258	0	0	51	107
$p < 0.01$	0	37	0	70	0	0	187	199	0	0	14	0
$p < 0.001$	0	0	0	31	0	0	146	131	0	0	0	0
σ	201.7	244.3	272.6	578.2	212.9	146.9	147.4	188.9	272.6	171.0	166.1	498.2

Note: Weighted results; unrestricted analysis sample. $\hat{\tau}$ represents the assumed or estimated rank correlation coefficient between potential treatment and control outcomes. Gains and losses: 'Total' refers to cumulative gains and losses, irrespective of their significance levels. Gains and losses denoted by $p < 0.05$, $p < 0.01$ and $p < 0.001$ are statistically significant at the 5%, 1% and 0.1% levels, respectively. σ denotes the standard deviation of (generalized) quantile treatment effects. Bootstrap standard errors (100 replications) were clustered to account for repeated observations.

References

Abadie, A. (2002): "Bootstrap Tests of Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284–292.

Abbring, J. and J. J. Heckman (2007): "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, Elsevier, Amsterdam, vol. 6B, chap. 72, 5145–5303.

Ashenfelter, O. (1983): "Determining Participation in Income-Tested Social Programs," *Journal of the American Statistical Association*, 78, 517–525.

Athey, S. and G. Imbens (2016): "Recursive Partitioning for Heterogeneous Causal Effects," *Proceedings of the National Academy of Sciences (PNAS)*, 113, 7353–7360.

Bitler, M., J. Gelbach, and H. Hoynes (2005): "Distributional impacts of the self-sufficiency project," Working Paper 11626, NBER, Cambridge, MA.

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006): "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, 96, 988–1012.

——— (2017): "Can Variation in Subgroups' Average Treatment Effects Explain Treatment Effect Heterogeneity? Evidence from a Social Experiment," *Review of Economics and Statistics*, 99, 683–697.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernández-Val (2018): "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments," NBER Working Paper Series 24678, National Bureau of Economic Research (NBER), Cambridge, MA.

Chernozhukov, V. and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2008): "Nonparametric Tests for Treatment Effect Heterogeneity," *Review of Economics and Statistics*, 90, 389–405.

Doksum, K. (1974): "Empirical Probability Plots and Statistical Inference for Nonlinear Models," *The Annals of Statistics*, 2, 267–277.

Fan, Y. and S. Park (2010): "Sharp Bounds on the Distribution of Treatment Effects and their Statistical Inference," *Econometric Theory*, 26, 931–951.

Fan, Y. and S. S. Park (2009): "Partial Identification of the Distribution of Treatment Effects and its Confidence Sets," *Advances inEconometrics*, 25, 3–70.

Fan, Y. and J. Wu (2010): "Partial Identification of the Distribution of Treatment Effects in Switching Regime Models and Its Confidence Sets," *The Review of Economic Studies*, 77, 1002–1041.

Firpo, S. (2007): "Efficient Semiparametric Estimation of Quantile Treatment Effects," *Econometrica*, 75, 259–276.

Firpo, S. and G. Ridder (2019): "Partial Identification of the Treatment Effect Distribution and its Functionals," *Journal of Econometrics*, 213, 210–234.

Frandsen, B. R. and L. J. Lefgren (2021): "Partial Identification of the Distribution of Treatment Effects with an Application to the Knowledge is Power Program (KIPP)," *Quantitative Economics*, 12.

Frank, M. J., R. B. Nelsen, and B. Schweizer (1987): "Best-possible Bounds for the Distribution of a Sum - A Problem of Kolmogorov," *Probability Theory and Related Fields*, 74, 199–211.

Frölich, M. and B. Melly (2013): "Unconditional Quantile Treatment Effects Under Endogeneity," *Journal of Business and Economic Statistics*, 31, 346–357.

Gibbons, J. D. and S. Chakraborti (2011): *Nonparametric Statistical Inference*, Boca Raton, FL: Chapman and Hall/CRC, 5th ed.

Gillitzer, C. and M. Sinning (2020): "Nudging Businesses to Pay their Taxes: Does Timing Matter?" *Journal of Economic Behavior and Organization*, 169, 284–300.

Hahn, P. R., J. S. Murray, and C. Carvalho (2020): "Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding, and Heterogeneous Effects," *Bayesian Analysis*, 15, 965–1056.

Heckman, J. J., H. Ichimura, and P. Todd (1998): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65, 261–294.

Heckman, J. J. and R. Robb (1985): "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, ed. by J. J. Heckman and B. Singer, Cambridge University Press, vol. 10, 156–245.

Heckman, J. J., J. Smith, and N. Clements (1997): "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts," *Review of Economic Studies*, 64, 487–535.

Heckman, J. J. and E. J. Vytlacil (2007): "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Economic Estimators to Evaluate Social Programs and to Forecast Their Effects in New Environments," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, Elsevier, Amsterdam, vol. 6B, chap. 71, 4875–5143.

Hill, J. (2011): "Bayesian Nonparametric Modeling for Causal Inference," *Journal of Computational and Graphical Statistics*, 20, 217–240.

Imai, K. and M. Ratkovic (2013): "Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation," *The Annals of Applied Statistics*, 7, 443–470.

Imbens, G. W. and D. B. Rubin (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge Books, Cambridge University Press.

Kendall, M. G. (1938): "A New Measure of Rank Correlation," *Biometrika*, 30, 81–93.

Kendall, M. G. and J. D. Gibbons (1990): *Rank Correlation Methods*, New York: Oxford University Press, 5th ed.

Kim, J. H. (2018): "Identifying the Distribution of Treatment Effects under Support Restrictions," arXiv, e-print 1410.5885.

Kline, P. and M. Tartari (2016): "Bounding the Labor Supply Response to a Randomized Welfare Experiment: A Revealed Preference Approach," *American Economic Review*, 106, 972–1014.

Künzel, S. R., J. S. Sekhon, P. J. Bickel, and B. Yu (2019): "Metalearners for Estimating Heterogeneous Treatment Effects using Machine Learning," *Proceedings of the National Academy of Sciences (PNAS)*, 116, 4156–4165.

Lechner, M. (1999): "Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification," *Journal of Business and Economic Statistics*, 17, 74–90.

Lee, M. J. (2006): *Micro-Econometrics for Policy, Program and Treatment Effects*, Oxford University Press.

Lee, S. (2022): "Partial Identification and Inference for Conditional Distributions of Treatment Effects," arXiv, e-print 2108.00723.

Lee, S. and A. M. Shaikh (2013): "Multiple Testing and Heterogeneous Treatment Effects: Re-evaluating the Effects of Progresa on School Enrollment," *Journal of Applied Econometrics*, 29, 612–626.

Lehmann, E. (1974): *Nonparametrics: Statistical Methods Based on Ranks*, San Francisco: Holden-Day.

List, J. A., A. M. Shaikh, and Y. Xu (2019): "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, 22, 773–793.

Makarov, G. D. (1982): "Estimates for the Distribution Function of a Sum of Two Random Variables When the Marginal Distributions are Fixed," *Theory of Probability and its Applications*, 26, 803–806.

Manski, C. (1997): "The Mixing Problem in Programme Evaluation," *Review of Economic Studies*, 64, 537–553.

Nie, X. and S. Wager (2021): "Quasi-Oracle Estimation of Heterogeneous Treatment Effects," *Biometrika*, 108, 299–319.

Parzen, E. (1979): "Nonparametric Statistical Data Modeling," *Journal of the American Statistical Association*, 74, 105–121.

Powers, S., J. Qian, K. Jung, A. S. N. H. Shah, T. Hastie, and R. Tibshirani (2018): "Some Methods for Heterogeneous Treatment Effect Estimation in High Dimensions," *Statistics in Medicine*, 37, 1767–1787.

Rosenbaum, P. and D. B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55.

Rubin, D. (1990): "Formal Modes of Statistical Inference for Causal Effects," *Journal of Statistical Planning and Inference*, 25, 279–292.

Rubin, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies," *Journal of Educational Psychology*, 66, 688–701.

Rüschendorf, L. (1982): "Random Variables with Maximum Sums," *Advances in Applied Probability*, 14, 623–632.

Russel, T. M. (2021): "Sharp Bounds on Functionals of the Joint Distribution in the Analysis of Treatment Effects," *Journal of Business and Economic Statistics*, 39, 532–546.

Shalit, U., F. D. Johansson, and D. Sontag (2017): "Estimating Individual Treatment Effect: Generalization Bounds and Algorithms," *Proceedings of Machine Learning Research*, 3076–3085.

Su, X., C.-L. Tsai, H. Wang, D. M. Nickerson, and B. Li (2009): "Subgroup Analysis via Recursive Partitioning," *Journal of Machine Learning Research*, 10, 141–158.

Wager, S. and S. Athey (2018): "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113, 1228–1242.

Williamson, R. C. and T. Downs (1990): "Probabilistic Arithmetic I. Numerical Methods for Calculating Convolutions and Dependency Bounds," *International Journal of Approximate Reasoning*, 4, 89–158.

Appendix A – Additional results

1. QTE results under rank invariance

Figure A1 displays the replication results of conventional QTE estimates originally presented in Bitler et al. (2006), using the approach of Firpo (2007). In contrast to Bitler et al. (2006), who report 90 percent confidence intervals, we present 95 percent confidence intervals throughout the paper. While we were unable to replicate the exact confidence intervals of Bitler et al. (2006) due to their reliance on bootstrap standard errors, we adopt their approach of estimating clustered bootstrap standard errors to account for the presence of repeated observations.

2. Intuition behind the identification of τ_{sp}

To provide context for the identification of τ_{sp}, it is useful to examine the values of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ and $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ we would expect to observe if τ_{sp} was known. Table A1 presents simulated values of $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ corresponding to various values of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ and τ_{sp} based on averaging over 100 random permutations of 2 × 16,772 observation units ($N_1 = N_0 = N/2$ =16,772), which matches our restricted analysis sample. For simplicity, we assume no ties and consider a model with a single continuous covariate vector $\mathbf{X}_1$ of length $N \times 1$. We choose 100 random permutations of $\mathbf{X}_1$ for different values of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) = \tau(\mathbf{Y}_0, \mathbf{X}_1)$. We also choose 100 random permutations of $\mathbf{Y}_1$ for various values of τ_{sp}. We can use the permutations of $\mathbf{Y}_1$ to obtain simulated values of $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1) = \tau(\mathbf{Y}_1, \mathbf{X}_1)$ because the ranks of $\mathbf{X}_1$ in the treatment group sample are the same as those in the control group sample.

We initially consider the case in which the rank invariance assumption holds, $\tau_{\text{sp}} = 1$. In this case, we expect to observe an identical rank correlation coefficient between actual and predicted outcomes in both groups, $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1) = \tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$, irrespective of how well we predict the ranks of the control outcomes. The values in Column (1) of Table A1 confirm this expectation: $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ equals $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ if $\tau_{\text{sp}} = 1$. Similarly, for cases in which the set of covariates predicts the ranks of the control outcomes perfectly, $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) = 1$, any deviation of $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ from 1 is expected to emanate from an imperfect rank correlation between potential treatment and control outcomes.

The first row of Table A1 validates this point by showing that $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ equals τ_{sp} if $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) = 1$.

We observe that for the remaining cases in which $0 \leq \tau_{\text{sp}} < 1$ and $0 \leq \tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) < 1$, $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ deviates from τ_{sp} because our predictions of the ranks of control outcomes are not perfect. At the same time, $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ deviates from $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ because the rank invariance assumption does not hold. The simulated values presented in Table A1 form a symmetric matrix because deviations of τ_{sp} from 1 have the same impact on $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ as deviations of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ from 1. Table A1 focuses on cases in which τ_{sp} and $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ are non-negative. Extending the analysis to cases in which $\tau_{\text{sp}} < 0$ and $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) < 0$ produces a mirror image of the positive values presented in Table A1. Similarly, analyzing cases in which $\tau_{\text{sp}} < 0$ and $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) \geq 0$ or $\tau_{\text{sp}} \geq 0$ and $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0) < 0$ produces a negative mirror image of the positive values in Table A1.

3. Testing the predictive strength of covariates

We test the predictive strength of covariates by creating 100 simulated datasets. Each simulated dataset s, $s = \{1, \ldots, 100\}$, contains observed outcomes and predicted values, with permutations of observation units tailored to a specific target value τ_{sp} and a specific rank correlation coefficient $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$. We compare the target values to estimated rank correlation coefficients. For simplicity, we use our restricted analysis sample to create simulated datasets with equal treatment and control group sizes.[26] Within each dataset, we find a random permutation of $\mathbf{Y}_1$ that satisfies the RCP. We use the permutations derived from the 100 datasets to estimate $\widehat{\tau} = (1/100) \sum_{s=1}^{100} \tau(\mathbf{\Pi}_s \mathbf{Y}_1, \mathbf{Y}_0)$, where $\mathbf{\Pi}_s$, $s = \{1, \ldots, 100\}$, is a $(N_1 \times N_1)$-permutation matrix.

To ascertain whether our estimate deviates significantly from the target value, we conduct a hypothesis test with the null hypothesis $H_0 : \tau_{\text{sp}} - \widehat{\tau} = 0$ against the alternative hypothesis $H_1 : \tau_{\text{sp}} - \widehat{\tau} \neq 0$. Our test results, which are presented in Panel A of Table A2, confirm that our predictors generally possess sufficient strength to yield unbiased estimates of the rank correlation between potential treatment and control outcomes. However, we do observe a small but statistically significant bias for transfers during the pre-time limit period.

[26] It is possible to accommodate different treatment and control group sizes by using an approach akin to that employed in Section 4.4.2.

To explore how deviations in estimated rank correlation coefficients from target values affect our QDTE estimates, we obtain QDTE estimates corresponding to different values of τ_{sp}. We use a two-sided Kolmogorov-Smirnov test to determine whether differences between alternative rank correlation coefficients translate into significant differences between QDTE functions. The test results, which are presented in Panel B of Table A2, indicate that the differences between estimated rank correlation coefficients and target values have no significant impact on our QDTE estimates.

4. Implications of assuming equally likely permutations

The primary identifying assumption in this paper is that all permutations satisfying the RCP are equally likely. While it is not possible to test this assumption, we can explore its implications for GQTE estimates, which are derived from averages of permutation-specific GQTE estimates. In Figure A2, we present permutation-specific GQTEs for income during the pre-time limit period. Figures A2a-A2e illustrate the density functions of permutation-specific GQTEs at selected quantiles. Our findings indicate that these density functions converge toward approximately normal distributions as the number of permutations increases. Moreover, we observe that even a relatively modest number of permutations ($P = 100$) yields a reasonably accurate approximation.

Figure A2f depicts the GQTE and the underlying density functions of permutation-specific GQTEs across percentiles. We observe that 50 percent of the probability mass of the underlying density functions is concentrated in close proximity to the GQTE. This observation provides strong evidence in support of calculating averages under the assumption of equally likely permutations. Our findings also emphasize the limitations on knowledge generation when deriving bounds from permutations with extremely low probability of occurrence. Moreover, our results highlight the potential for developing tests to compare alternative identifying assumptions. For instance, one could compare GQTE based on equally likely permutations to GQTE based on an alternative weighting scheme. The exploration of this issue remains a subject for future research.

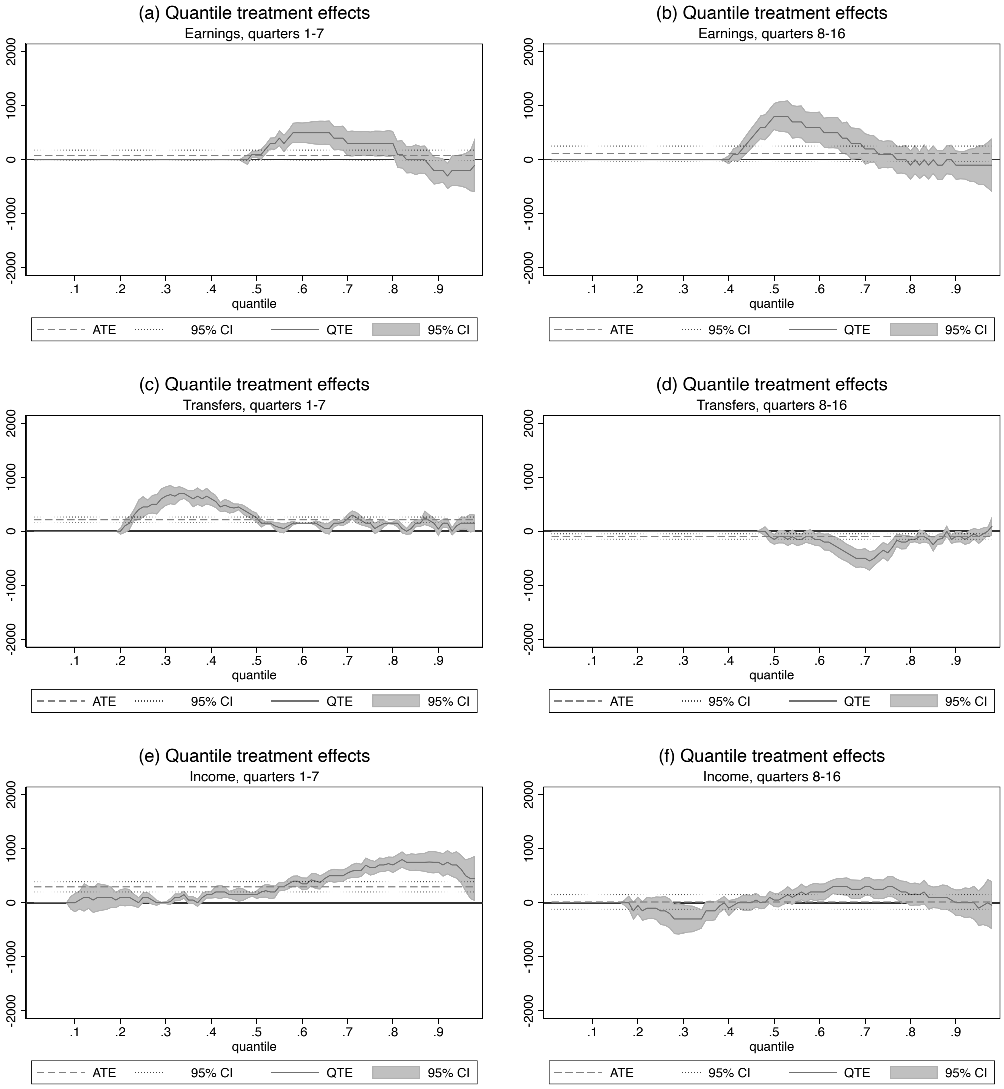
5. Impact of sample restriction and reweighting

Figure A3 presents QDTE, GQTE, and QTE estimates for income during the pre-time limit period, illustrating the consequences of two factors: imposing a sample restriction on the AFDC sample and employing a reweighting approach to control for covariates. We find that excluding 11 randomly selected observations from the AFDC sample to obtain a restricted analysis sample with balanced treatment and control group sizes has minimal influence on our results. We also observe that, due to random assignment of observation units to treatment and control groups, the use of a reweighting approach does not change our results qualitatively.

Figures A3c and A3d depict unweighted and weighted QTE and GQTE estimates under rank invariance for our unrestricted analysis sample.[27] These results confirm that our approach yields very similar results to the approach of Firpo (2007) under rank invariance. While the QTE estimates presented in Figure A3d involve calculating reweighted quantiles of unweighted outcomes, our GQTE estimates are based on calculating unweighted quantiles of reweighted outcomes. We have no clear preference for one approach over another if the rank invariance assumption holds. However, in contrast to the approach of Firpo (2007), our approach remains applicable in situations where the rank invariance assumption is violated.

[27] Unweighted QTE and GQTE estimates under rank invariance for our restricted analysis sample are presented in Figure 2f.

Figure A1: Quantile Treatment Effects under Rank Invariance



Note: Weighted results; unrestricted analysis sample. ATE: Confidence intervals are based on clustered standard errors to account for repeated observations. QTE: Confidence intervals are based on clustered bootstrap standard errors (100 replications) to account for repeated observations. Top percentile not included due to high sampling variability.

Table A1: Simulated values of $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ for given values of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ and τ_{sp}

	τ_{sp}										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.0
1.0	1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00
0.9	0.90	0.86	0.78	0.69	0.59	0.49	0.39	0.30	0.20	0.10	0.00
0.8	0.80	0.78	0.72	0.64	0.56	0.47	0.38	0.28	0.19	0.10	0.00
0.7	0.70	0.69	0.64	0.59	0.51	0.43	0.35	0.26	0.18	0.09	0.00
0.6	0.60	0.59	0.56	0.51	0.46	0.39	0.32	0.24	0.16	0.08	0.00
0.5	0.50	0.49	0.47	0.43	0.39	0.33	0.27	0.21	0.14	0.07	0.00
0.4	0.40	0.39	0.38	0.35	0.32	0.27	0.22	0.17	0.12	0.06	0.00
0.3	0.30	0.30	0.28	0.26	0.24	0.21	0.17	0.13	0.09	0.05	0.00
0.2	0.20	0.20	0.19	0.18	0.16	0.14	0.12	0.09	0.06	0.03	0.00
0.1	0.10	0.10	0.09	0.09	0.08	0.07	0.06	0.04	0.03	0.02	-0.00
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00

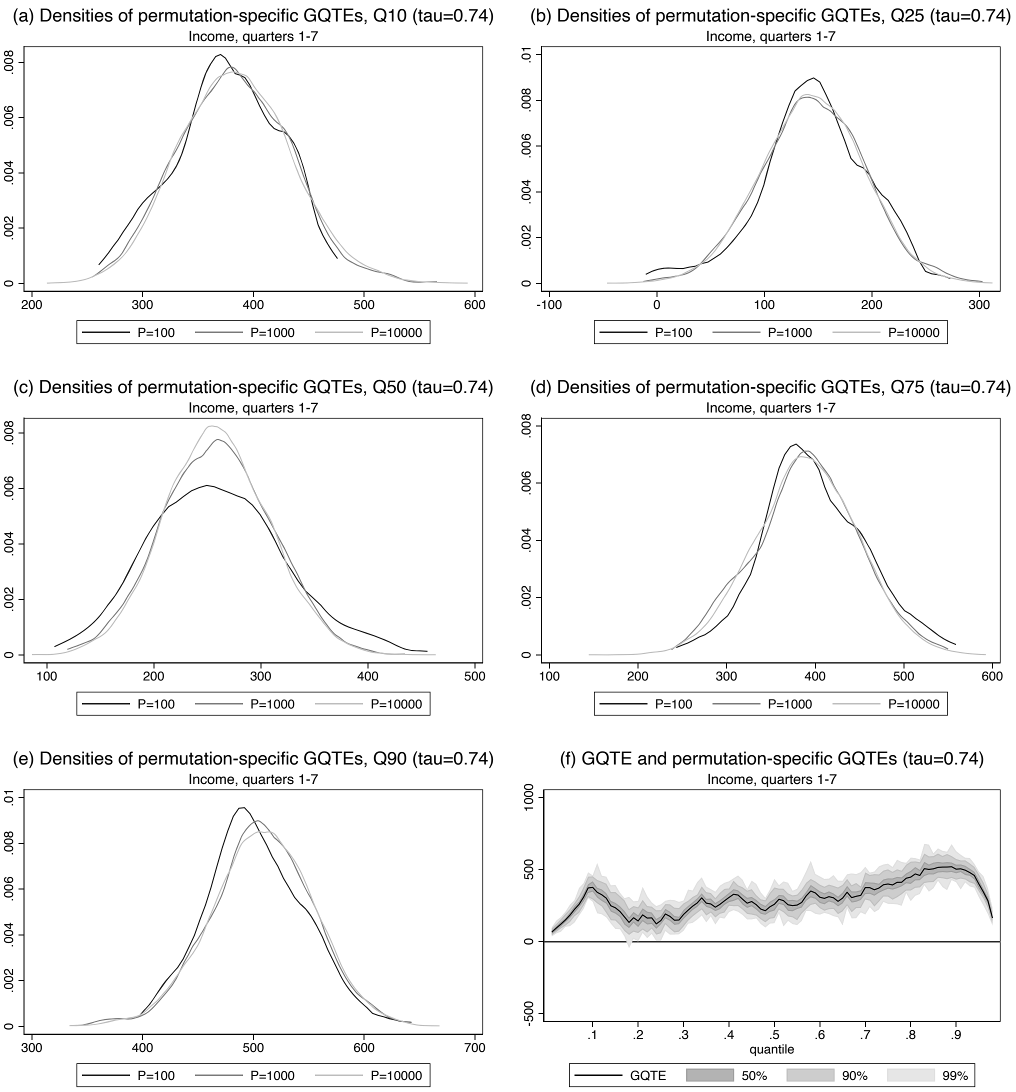
Note: This table presents simulated values of $\tau(\mathbf{Y}_1, \widehat{\mathbf{Y}}_1)$ using simulation data for alternative values of $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$ and τ_{sp}. Each simulated value is based on 100 random permutations of 2 × 16,772 observation units.

Table A2: Simulation-based test for strength of model predictors and Kolmogorov-Smirnov test for equality of distribution functions

	Earnings		Transfers		Income	
	(1)	(2)	(3)	(4)	(5)	(6)
	Q1-7	Q8-16	Q1-7	Q8-16	Q1-7	Q8-16
Panel A.						
$\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$	0.35	0.31	0.35	0.30	0.32	0.27
τ_{sp}	Estimate					
0.80			0.78 [0.00]			
0.76					0.74 [0.06]	
0.72	0.71 [0.46]					
0.70				0.69 [0.17]		
0.61		0.62 [0.29]				
0.60						0.61 [0.20]
Panel B.						
τ_1, τ_2	Kolmogorov-Smirnov test (p-values)					
0.80, 0.78			0.994			
0.76, 0.74					0.696	
0.72, 0.71	1.000					
0.70, 0.69				1.000		
0.61, 0.62		0.577				
0.60, 0.61						1.000

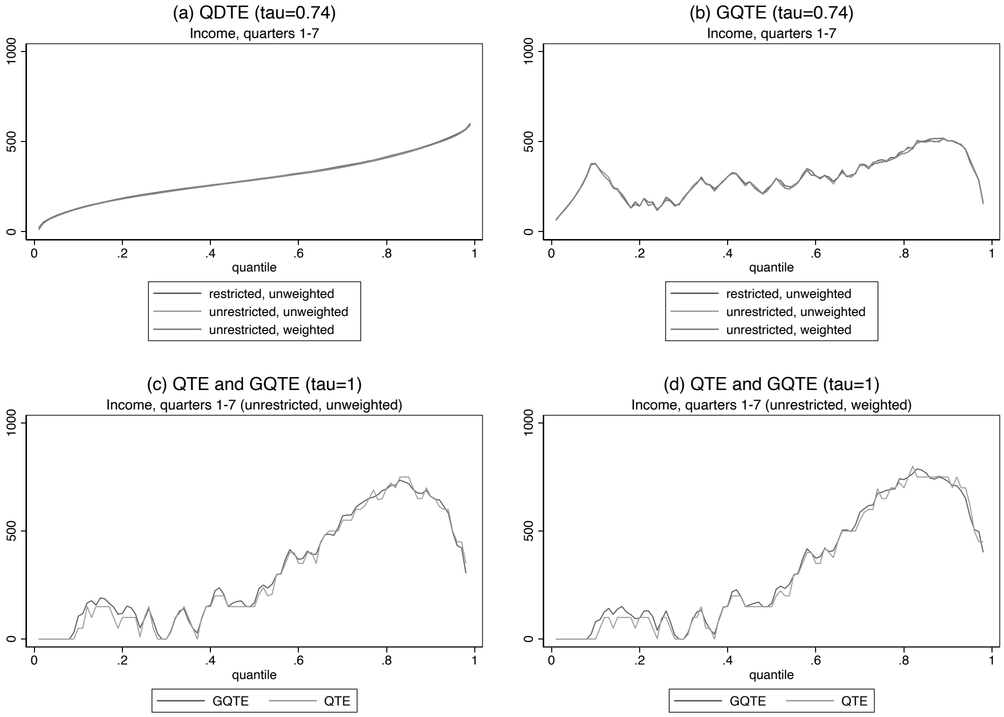
Note: Panel A: We generate 100 simulation datasets based on our restricted analysis sample. Each dataset contains random permutations of observation units consistent with a specific target value τ_{sp} and a given rank correlation between actual and predicted control outcomes $\tau(\mathbf{Y}_0, \widehat{\mathbf{Y}}_0)$. We use these datasets to estimate $\widehat{\tau}$ while varying τ_{sp} over the range from 0.6 to 0.8. Panel A presents the target values and their corresponding estimates, which conform with the estimates shown in Figures 1a and 1b. We test whether the target values deviate significantly from the corresponding estimates. The associated p-values for the two-sided test comparing target values and estimates are reported in brackets. Panel B: Two-sided Kolmogorov-Smirnov test. We compare QDTE estimates under two different assumed values of the rank correlation coefficient between potential treatment and control outcomes, τ_1 and τ_2. Panel B presents the associated p-values, which were calculated using a counting algorithm as described in Gibbons and Chakraborti (2011).

Figure A2: GQTE and permutation-specific GQTEs,
Income, quarters 1-7



Note: GQTE estimates are derived from averaging permutation-specific GQTE estimates. Figures A2a-A2e: density functions of permutation-specific GQTEs at selected quantiles. Figure A2f: GQTE and underlying density functions of permutation-specific GQTEs across percentiles. Top percentile not included due to high sampling variability.

Figure A3: Impact of sample restriction and reweighting, Income, quarters 1-7



Note: Figures A3b-A3d: Top percentile not included due to high sampling variability.

Appendix B – Monte Carlo simulation

We use Monte Carlo simulations to study the finite sample behavior of our estimators. We consider two normally distributed $(N \times 1)$-vectors, $\mathbf{Y}_1 \sim N(\mu_1, \sigma_1^2)$ and $\mathbf{Y}_0 \sim N(\mu_0, \sigma_0^2)$. Moreover, we consider a single $(N \times 1)$-covariate vector $\mathbf{X}_1$, which contains the ranks of a random variable without ties.

1. Rank correlation coefficients

We use random permutations of $\mathbf{Y}_1$ to generate simulation datasets in which the rank correlation coefficient $\tau_{\text{sp}} = \tau(\mathbf{Y}_1, \mathbf{Y}_0)$ ranges from 0 to 0.9. We also use random permutations to vary the predictive strength of $\mathbf{X}_1$ by adjusting the rank correlation coefficient $\tau_c = \tau(\mathbf{Y}_0, \mathbf{X}_1)$ over the range from 0.1 to 0.9. For each combination of τ_{sp} and τ_c, we create 100 simulation datasets with a sample size of 500 (250 observation units in each group) and an additional 100 simulation datasets with a sample size of 5,000 (2,500 observation units in each group). Within each simulation dataset, we estimate τ_{sp} using 100 random permutations of $\mathbf{Y}_1$ satisfying the RCP. We obtain bootstrap standard errors for each estimate using 100 replications. We repeat this process 100 times for each sample and each combination of τ_{sp} and τ_c to obtain our Monte Carlo simulation results.

Tables B1 and B2 present the results for datasets with 500 and 5,000 observation units, respectively. The results confirm the precision of our estimator of τ_{sp}, even when the sample size is relatively small. They also confirm that the estimator is unbiased when τ_c exceeds a certain threshold. Specifically, for $\tau_c \geq 0.4$, our estimator performs well in terms of bias and root mean square error (RMSE). For $\tau_c \geq 0.3$, our estimator maintains its accuracy as long as τ_{sp} remains below or equal to 0.8. We also find that the biases diminish notably as the sample size increases. Tables B3 and B4 report the estimated bootstrap standard errors, their lower and upper 5th percentile, and the 90 percent coverage rate. The results confirm that our estimated bootstrap standard errors are a good representation of the true sampling variation.

2. QID, QDTE and GQTE

We study the finite sample behavior of QID, QDTE, and GQTE estimators. We derive target values for the case in which $\mathbf{Y}_1$ and $\mathbf{Y}_0$ are independent to establish their statistical properties, including unbiasedness, consistency, and asymptotic normality. We also assess the validity of estimated bootstrap standard errors for the QDTE and GQTE estimators.

For $u \in (0,1)$, the quantile function of the QID estimator under independence is given by

$$(\mu_1 - \mu_0) + \sqrt{\sigma_1^2 + \sigma_0^2}\Phi^{-1}(u),$$

where $\Phi(\cdot)$ represents the CDF of the standard normal distribution. Following Fan and Park (2009), we consider a scenario where $\mathbf{Y}_1 \sim N(2,2)$ and $\mathbf{Y}_0 \sim N(1,1)$. We estimate the QID at selected quantiles, including Q10, Q25, Q50, Q75, and Q90, and subsequently compare our estimates to the target values of the quantile function at these quantiles. The respective target values are −1.22, −0.17, 1.00, 2.17, and 3.22.

The quantile function of the QID estimator is based on arranging treatment effects in ascending order, capturing the full amount of heterogeneity in the data, without considering the location of control outcomes. Our objective is to compare each control outcome to a set of treatment outcomes resulting from permutations that have a positive probability of occurrence. Under independence, this means we compare each control outcome to all possible treatment outcomes. Assuming equally likely permutations, we calculate the average over all permutations to obtain the GQTE. Therefore, the quantile function of the GQTE estimator under independence is obtained by comparing the expected value of the treatment outcome, μ_1, to the quantile function of $\mathbf{Y}_0$. The quantile function of $\mathbf{Y}_0$ is given by $\mu_0 + \sigma_0\Phi^{-1}(u)$. Consequently, the quantile function of the GQTE estimator under independence is

$$\mu_1 - (\mu_0 + \sigma_0\Phi^{-1}(u)).$$

In the scenario where $\mathbf{Y}_1 \sim N(2,2)$ and $\mathbf{Y}_0 \sim N(1,1)$, the target values at Q10, Q25, Q50, Q75, and Q90 are 2.28, 1.67, 1.00, 0.33, and −0.28, respectively. To derive the quantile function of the QDTE estimator, the values of the quantile function of the GQTE estimator have to be rearranged to be monotonically increasing. Therefore, the corresponding target values of the QDTE are −0.28, 0.33, 1.00, 1.67, and 2.28.

Figures B1 and B2 provide evidence of the unbiasedness, consistency, and asymptotic normality of our estimators. We focus on a scenario in which $\tau_c = 0.4$. Estimation is based on 100 random permutations of observation units. Figure B1 illustrates the convergence of our estimators towards the relevant target values at sample sizes of 500, 5,000, and 50,000. Figure B2 illustrates that the distributions of a recentered and rescaled version of our estimators become increasingly indistinguishable from a normal distribution as the sample size increases. This finding underscores the asymptotic normality property of our estimators. Figures B3 and B5 present bootstrap standard errors for selected quantiles of QDTE and GQTE estimators. These findings confirm that increased sample size and the availability of highly predictive covariates contribute to enhanced precision. In Figures B4 and B6, we report the 90 percent coverage rates of QDTE and GQTE estimators. These results validate the accuracy of our estimated bootstrap standard errors.

Table B1: Monte Carlo simulation (500 observations; 100 replications): Point estimates of τ_{sp} resulting from 100 random permutations

τ_{sp}	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00
τ_c =.9										
Mean	0.899	0.800	0.699	0.599	0.500	0.400	0.301	0.200	0.099	0.010
Standard deviation	0.008	0.008	0.008	0.007	0.007	0.008	0.009	0.008	0.008	0.001
Bias	-0.001	0.000	-0.001	-0.001	0.000	0.000	0.001	-0.000	-0.001	0.010
RMSE	0.008	0.008	0.008	0.007	0.007	0.008	0.009	0.008	0.008	0.010
τ_c =.8										
Mean	0.909	0.801	0.699	0.602	0.498	0.402	0.301	0.199	0.101	0.013
Standard deviation	0.016	0.016	0.014	0.014	0.014	0.014	0.016	0.015	0.015	0.006
Bias	0.009	0.001	-0.001	0.002	-0.002	0.002	0.001	-0.001	0.001	0.013
RMSE	0.018	0.016	0.014	0.014	0.014	0.015	0.016	0.015	0.015	0.015
τ_c =.7										
Mean	0.912	0.801	0.699	0.604	0.501	0.403	0.297	0.203	0.101	0.018
Standard deviation	0.020	0.027	0.023	0.022	0.019	0.022	0.021	0.023	0.025	0.011
Bias	0.012	0.001	-0.001	0.004	0.001	0.003	-0.003	0.003	0.001	0.018
RMSE	0.023	0.027	0.023	0.023	0.019	0.022	0.022	0.023	0.025	0.021
τ_c =.6										
Mean	0.892	0.801	0.697	0.604	0.498	0.401	0.297	0.200	0.105	0.021
Standard deviation	0.023	0.036	0.031	0.034	0.031	0.029	0.033	0.030	0.036	0.014
Bias	-0.008	0.001	-0.003	0.004	-0.002	0.001	-0.003	-0.000	0.005	0.021
RMSE	0.024	0.036	0.031	0.035	0.031	0.029	0.033	0.030	0.036	0.026
τ_c =.5										
Mean	0.873	0.800	0.699	0.612	0.495	0.405	0.310	0.206	0.109	0.031
Standard deviation	0.023	0.037	0.047	0.045	0.047	0.046	0.045	0.048	0.040	0.025
Bias	-0.027	0.000	-0.001	0.012	-0.005	0.005	0.010	0.006	0.009	0.031
RMSE	0.036	0.037	0.047	0.046	0.048	0.047	0.046	0.048	0.041	0.040
τ_c =.4										
Mean	0.843	0.777	0.698	0.627	0.505	0.413	0.306	0.196	0.106	0.035
Standard deviation	0.023	0.044	0.043	0.069	0.065	0.058	0.061	0.062	0.057	0.033
Bias	-0.057	-0.023	-0.002	0.027	0.005	0.013	0.006	-0.004	0.006	0.035
RMSE	0.062	0.049	0.043	0.075	0.065	0.060	0.061	0.062	0.058	0.048
τ_c =.3										
Mean	0.813	0.758	0.704	0.616	0.509	0.410	0.320	0.217	0.116	0.058
Standard deviation	0.027	0.046	0.061	0.097	0.082	0.094	0.095	0.086	0.070	0.045
Bias	-0.087	-0.042	0.004	0.016	0.009	0.010	0.020	0.017	0.016	0.058
RMSE	0.091	0.062	0.061	0.098	0.083	0.094	0.097	0.088	0.072	0.073
τ_c =.2										
Mean	0.772	0.728	0.671	0.615	0.524	0.438	0.341	0.235	0.155	0.104
Standard deviation	0.024	0.050	0.076	0.092	0.119	0.116	0.129	0.125	0.122	0.094
Bias	-0.128	-0.072	-0.029	0.015	0.024	0.038	0.041	0.035	0.055	0.104
RMSE	0.130	0.087	0.081	0.093	0.122	0.122	0.136	0.129	0.134	0.140
τ_c =.1										
Mean	0.725	0.678	0.655	0.589	0.551	0.481	0.371	0.351	0.282	0.173
Standard deviation	0.024	0.056	0.070	0.109	0.130	0.170	0.210	0.197	0.214	0.168
Bias	-0.175	-0.122	-0.045	-0.011	0.051	0.081	0.071	0.151	0.182	0.173
RMSE	0.177	0.134	0.083	0.110	0.140	0.189	0.221	0.248	0.281	0.241

Table B2: Monte Carlo simulation (5,000 observations; 100 replications):
Point estimates of τ_{sp} resulting from 100 random permutations

τ_{sp}	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00
τ_c =.9										
Mean	0.901	0.799	0.699	0.599	0.500	0.401	0.301	0.200	0.100	0.007
Standard deviation	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.001
Bias	0.001	-0.001	-0.001	-0.001	-0.000	0.001	0.001	0.000	-0.000	0.007
RMSE	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.007
τ_c =.8										
Mean	0.907	0.800	0.698	0.599	0.499	0.400	0.301	0.200	0.100	0.009
Standard deviation	0.004	0.006	0.005	0.005	0.004	0.006	0.005	0.005	0.005	0.001
Bias	0.007	0.000	-0.002	-0.001	-0.001	-0.000	0.001	0.000	0.000	0.009
RMSE	0.009	0.006	0.006	0.005	0.005	0.006	0.005	0.005	0.005	0.009
τ_c =.7										
Mean	0.915	0.803	0.698	0.601	0.501	0.399	0.300	0.199	0.100	0.010
Standard deviation	0.009	0.008	0.007	0.008	0.007	0.007	0.008	0.008	0.007	0.002
Bias	0.015	0.003	-0.002	0.001	0.001	-0.001	0.000	-0.001	0.000	0.010
RMSE	0.018	0.009	0.007	0.008	0.007	0.007	0.008	0.008	0.007	0.010
τ_c =.6										
Mean	0.898	0.804	0.698	0.602	0.503	0.400	0.300	0.200	0.103	0.012
Standard deviation	0.006	0.011	0.010	0.010	0.010	0.012	0.010	0.010	0.008	0.004
Bias	-0.002	0.004	-0.002	0.002	0.003	0.000	-0.000	-0.000	0.003	0.012
RMSE	0.007	0.012	0.011	0.010	0.011	0.012	0.010	0.010	0.009	0.013
τ_c =.5										
Mean	0.874	0.795	0.696	0.599	0.501	0.400	0.301	0.203	0.102	0.016
Standard deviation	0.010	0.013	0.011	0.013	0.012	0.013	0.014	0.013	0.012	0.008
Bias	-0.026	-0.005	-0.004	-0.001	0.001	0.000	0.001	0.003	0.002	0.016
RMSE	0.027	0.013	0.012	0.013	0.012	0.013	0.014	0.013	0.012	0.018
τ_c =.4										
Mean	0.858	0.782	0.697	0.605	0.500	0.401	0.304	0.205	0.108	0.019
Standard deviation	0.013	0.015	0.013	0.018	0.020	0.018	0.018	0.018	0.017	0.011
Bias	-0.042	-0.018	-0.003	0.005	0.000	0.001	0.004	0.005	0.008	0.019
RMSE	0.044	0.023	0.014	0.019	0.020	0.018	0.019	0.019	0.019	0.022
τ_c =.3										
Mean	0.834	0.774	0.694	0.616	0.511	0.412	0.311	0.212	0.114	0.026
Standard deviation	0.013	0.019	0.018	0.029	0.024	0.030	0.026	0.025	0.027	0.016
Bias	-0.066	-0.026	-0.006	0.016	0.011	0.012	0.011	0.012	0.014	0.026
RMSE	0.067	0.032	0.019	0.033	0.027	0.032	0.029	0.028	0.030	0.030
τ_c =.2										
Mean	0.793	0.747	0.693	0.628	0.533	0.438	0.331	0.226	0.131	0.041
Standard deviation	0.014	0.023	0.020	0.033	0.042	0.042	0.047	0.042	0.039	0.030
Bias	-0.107	-0.053	-0.007	0.028	0.033	0.038	0.031	0.026	0.031	0.041
RMSE	0.107	0.058	0.021	0.043	0.053	0.056	0.057	0.049	0.050	0.051
τ_c =.1										
Mean	0.729	0.710	0.693	0.658	0.591	0.509	0.400	0.300	0.210	0.116
Standard deviation	0.011	0.016	0.019	0.036	0.064	0.074	0.084	0.083	0.082	0.070
Bias	-0.171	-0.090	-0.007	0.058	0.091	0.109	0.100	0.100	0.110	0.116
RMSE	0.171	0.092	0.020	0.068	0.111	0.132	0.131	0.130	0.137	0.136

Table B3: Bootstrap standard errors from Monte Carlo simulation
(500 observations; 100 replications)

τ_{sp}	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00
τ_c =.9										
Mean	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.002
Upper 5th percentile	0.009	0.009	0.009	0.009	0.008	0.009	0.009	0.009	0.009	0.003
Lower 5th percentile	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.002
90% coverage rate (%)	94	88	92	91	95	91	88	85	90	97
τ_c =.8										
Mean	0.016	0.016	0.014	0.015	0.015	0.016	0.015	0.015	0.015	0.006
Upper 5th percentile	0.018	0.018	0.016	0.017	0.017	0.017	0.017	0.017	0.017	0.008
Lower 5th percentile	0.014	0.014	0.012	0.013	0.013	0.014	0.014	0.013	0.013	0.005
90% coverage rate (%)	89	88	90	91	94	93	87	90	92	94
τ_c =.7										
Mean	0.020	0.026	0.022	0.023	0.024	0.023	0.022	0.024	0.023	0.010
Upper 5th percentile	0.022	0.029	0.025	0.026	0.026	0.026	0.025	0.026	0.026	0.013
Lower 5th percentile	0.018	0.022	0.020	0.021	0.022	0.021	0.019	0.021	0.021	0.008
90% coverage rate (%)	93	89	89	92	97	91	93	91	87	90
τ_c =.6										
Mean	0.021	0.036	0.032	0.034	0.033	0.032	0.032	0.033	0.031	0.016
Upper 5th percentile	0.023	0.040	0.037	0.038	0.037	0.036	0.036	0.036	0.035	0.019
Lower 5th percentile	0.019	0.032	0.027	0.029	0.028	0.028	0.028	0.029	0.028	0.013
90% coverage rate (%)	86	91	87	90	93	94	89	93	89	95
τ_c =.5										
Mean	0.022	0.042	0.043	0.047	0.044	0.043	0.044	0.044	0.042	0.022
Upper 5th percentile	0.025	0.047	0.048	0.052	0.049	0.049	0.049	0.048	0.047	0.026
Lower 5th percentile	0.020	0.037	0.037	0.041	0.038	0.038	0.039	0.039	0.037	0.018
90% coverage rate (%)	91	96	89	92	88	84	90	86	90	88
τ_c =.4										
Mean	0.024	0.044	0.055	0.062	0.062	0.059	0.060	0.059	0.055	0.034
Upper 5th percentile	0.027	0.049	0.063	0.070	0.069	0.068	0.067	0.066	0.060	0.041
Lower 5th percentile	0.021	0.040	0.048	0.055	0.054	0.051	0.054	0.052	0.049	0.027
90% coverage rate (%)	91	88	96	86	91	90	92	88	90	92
τ_c =.3										
Mean	0.025	0.044	0.063	0.080	0.088	0.087	0.085	0.084	0.074	0.050
Upper 5th percentile	0.028	0.049	0.070	0.090	0.099	0.098	0.095	0.093	0.084	0.059
Lower 5th percentile	0.023	0.039	0.055	0.071	0.080	0.075	0.075	0.075	0.065	0.042
90% coverage rate (%)	85	88	88	84	91	89	88	88	96	94
τ_c =.2										
Mean	0.025	0.045	0.070	0.096	0.117	0.131	0.133	0.128	0.111	0.083
Upper 5th percentile	0.028	0.052	0.077	0.105	0.130	0.143	0.146	0.146	0.129	0.100
Lower 5th percentile	0.022	0.039	0.062	0.085	0.104	0.118	0.119	0.115	0.097	0.070
90% coverage rate (%)	95	87	92	89	89	95	90	94	89	90
τ_c =.1										
Mean	0.026	0.051	0.082	0.117	0.150	0.177	0.192	0.201	0.196	0.182
Upper 5th percentile	0.029	0.059	0.095	0.133	0.166	0.193	0.210	0.216	0.211	0.202
Lower 5th percentile	0.023	0.044	0.069	0.104	0.131	0.160	0.174	0.187	0.180	0.163
90% coverage rate (%)	93	92	95	95	94	96	87	93	90	92

Table B4: Bootstrap standard errors from Monte Carlo simulation (5,000 observations; 100 replications)

τ_{sp}	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00
τ_c =.9										
Mean	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.001
Upper 5th percentile	0.003	0.003	0.004	0.004	0.004	0.004	0.004	0.003	0.004	0.001
Lower 5th percentile	0.002	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.001
90% coverage rate (%)	88	90	91	90	95	92	92	94	92	94
τ_c =.8										
Mean	0.005	0.006	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.001
Upper 5th percentile	0.006	0.006	0.006	0.006	0.006	0.005	0.006	0.006	0.006	0.002
Lower 5th percentile	0.004	0.005	0.005	0.005	0.004	0.004	0.005	0.004	0.004	0.001
90% coverage rate (%)	90	92	92	91	96	84	90	91	93	95
τ_c =.7										
Mean	0.009	0.008	0.008	0.008	0.007	0.007	0.008	0.008	0.007	0.002
Upper 5th percentile	0.010	0.009	0.008	0.008	0.008	0.008	0.009	0.008	0.008	0.003
Lower 5th percentile	0.008	0.007	0.007	0.007	0.006	0.007	0.007	0.007	0.006	0.002
90% coverage rate (%)	88	88	93	87	92	91	87	91	88	94
τ_c =.6										
Mean	0.007	0.011	0.010	0.010	0.010	0.011	0.011	0.010	0.010	0.004
Upper 5th percentile	0.008	0.012	0.011	0.011	0.012	0.012	0.012	0.012	0.011	0.005
Lower 5th percentile	0.006	0.010	0.009	0.009	0.009	0.010	0.009	0.009	0.009	0.003
90% coverage rate (%)	91	93	88	90	92	86	89	90	93	94
τ_c =.5										
Mean	0.010	0.012	0.012	0.014	0.014	0.015	0.014	0.014	0.014	0.006
Upper 5th percentile	0.011	0.014	0.014	0.016	0.015	0.016	0.015	0.015	0.015	0.008
Lower 5th percentile	0.009	0.011	0.011	0.012	0.012	0.014	0.012	0.012	0.012	0.005
90% coverage rate (%)	92	91	95	91	93	91	89	90	93	87
τ_c =.4										
Mean	0.012	0.016	0.015	0.020	0.020	0.019	0.018	0.018	0.018	0.010
Upper 5th percentile	0.013	0.018	0.018	0.022	0.022	0.022	0.021	0.020	0.021	0.012
Lower 5th percentile	0.011	0.014	0.013	0.017	0.018	0.017	0.016	0.016	0.016	0.008
90% coverage rate (%)	85	89	91	93	88	90	90	91	95	90
τ_c =.3										
Mean	0.014	0.020	0.018	0.027	0.028	0.027	0.026	0.026	0.026	0.016
Upper 5th percentile	0.015	0.023	0.021	0.030	0.031	0.030	0.029	0.029	0.029	0.019
Lower 5th percentile	0.012	0.018	0.015	0.023	0.024	0.024	0.023	0.023	0.024	0.014
90% coverage rate (%)	90	93	88	83	92	90	90	92	93	90
τ_c =.2										
Mean	0.015	0.025	0.021	0.036	0.042	0.042	0.042	0.041	0.040	0.031
Upper 5th percentile	0.016	0.027	0.025	0.040	0.047	0.048	0.047	0.045	0.045	0.035
Lower 5th percentile	0.013	0.022	0.018	0.032	0.036	0.038	0.038	0.036	0.036	0.027
90% coverage rate (%)	91	97	96	96	92	90	86	89	88	90
τ_c =.1										
Mean	0.011	0.016	0.023	0.042	0.066	0.080	0.083	0.083	0.082	0.072
Upper 5th percentile	0.012	0.018	0.026	0.047	0.072	0.089	0.094	0.094	0.092	0.080
Lower 5th percentile	0.010	0.014	0.019	0.036	0.060	0.070	0.073	0.074	0.074	0.065
90% coverage rate (%)	82	95	95	93	91	92	88	92	89	93

Figure B1: Monte Carlo simulation (100 replications): Consistency

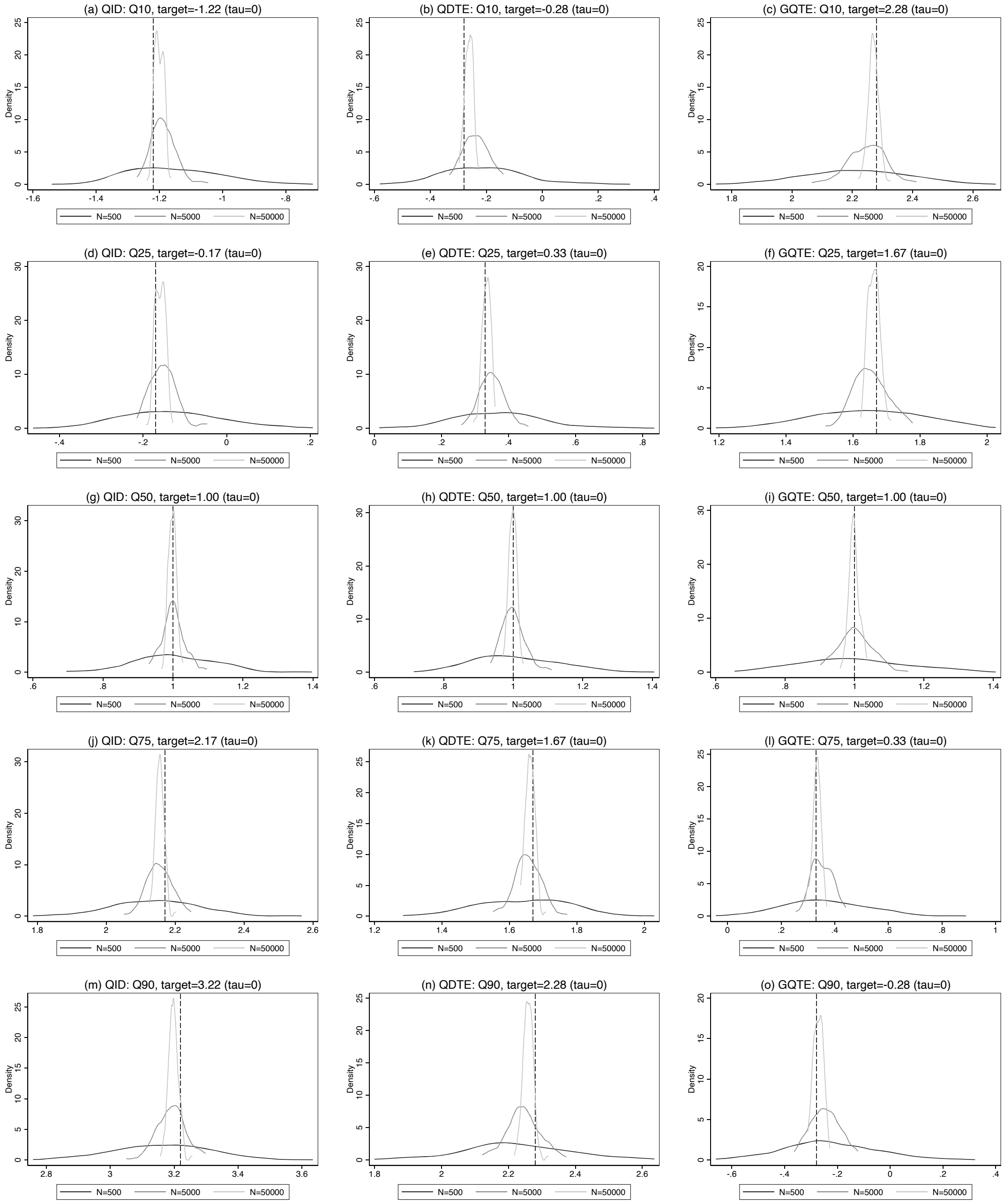


Figure B2: Monte Carlo simulation (100 replications): Asymptotic normality

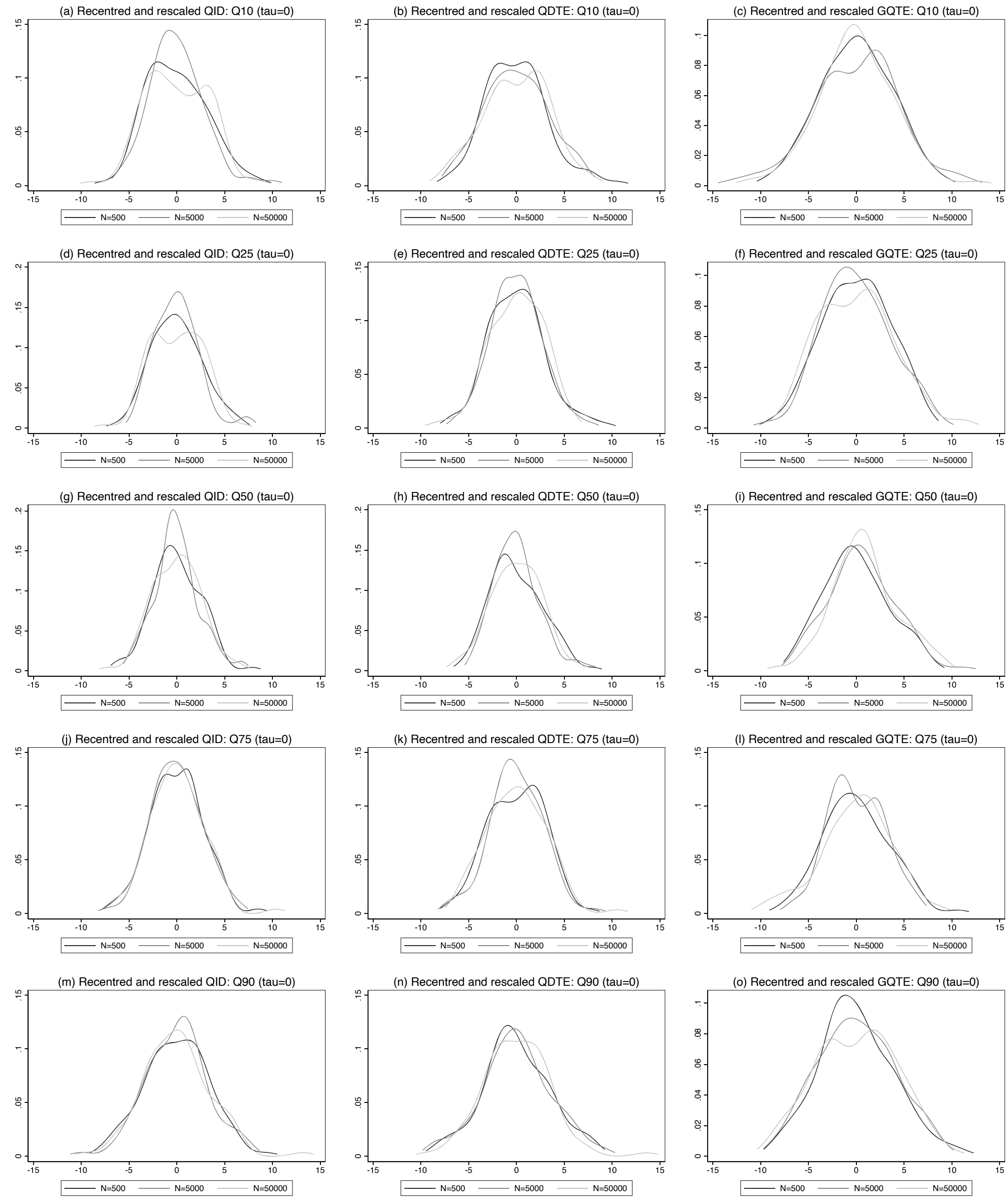


Figure B3: Monte Carlo simulation (100 reps.): Bootstrap standard errors for selected quantiles of QDTE

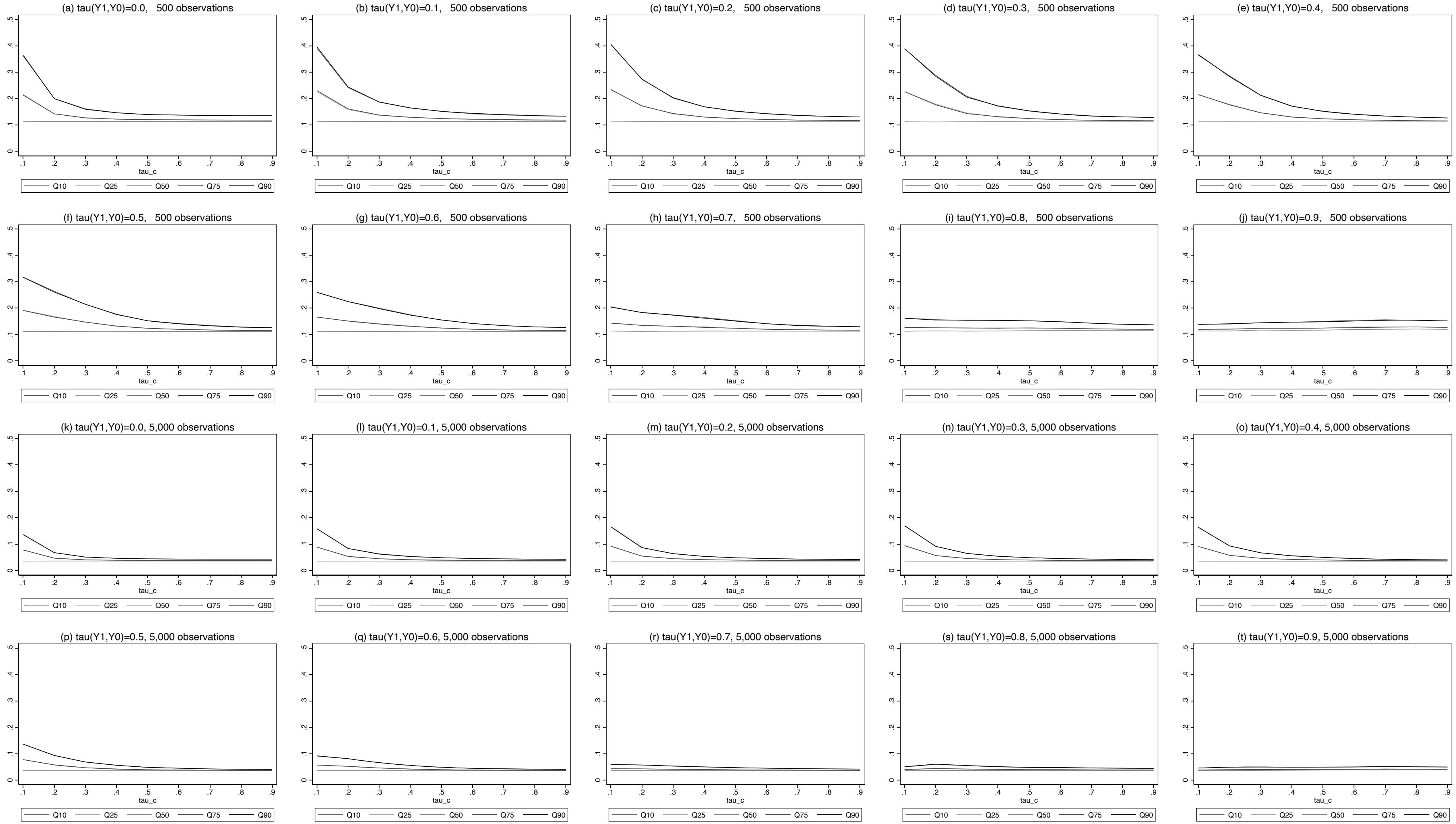


Figure B4: Monte Carlo simulation (100 replications): 90% coverage rate for selected quantiles of QDTE

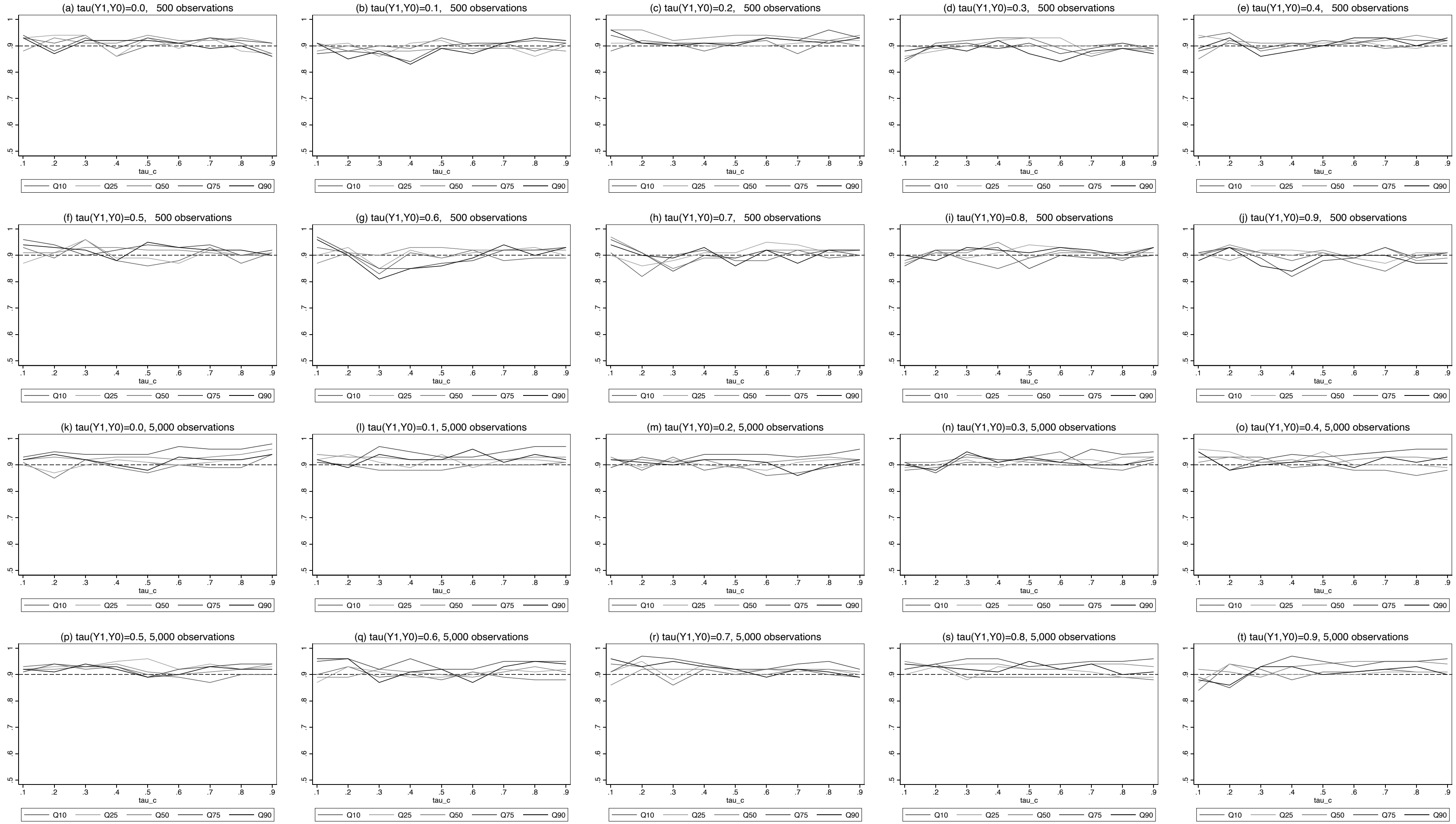


Figure B5: Monte Carlo simulation (100 reps.): Bootstrap standard errors for selected quantiles of GQTE

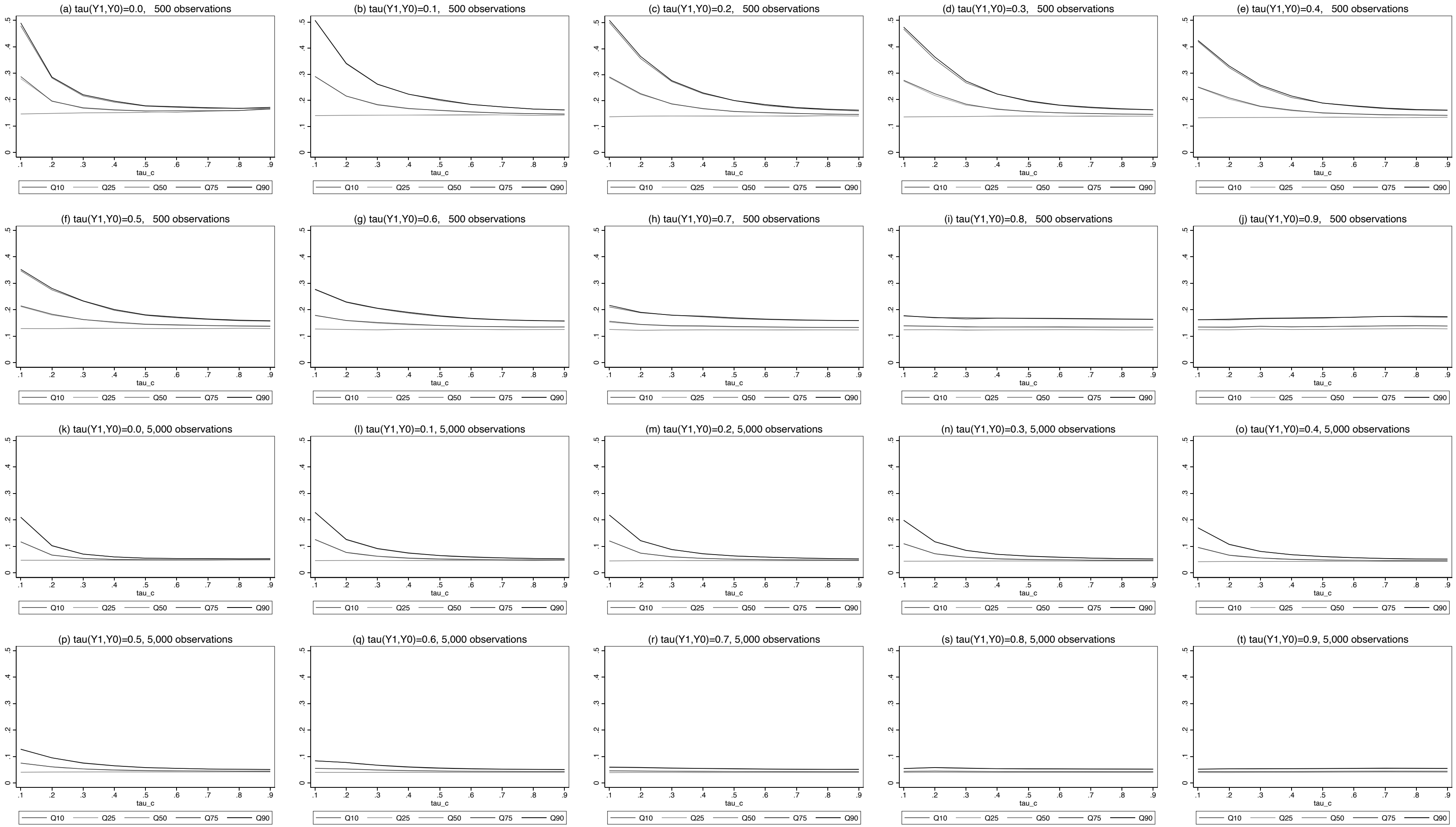
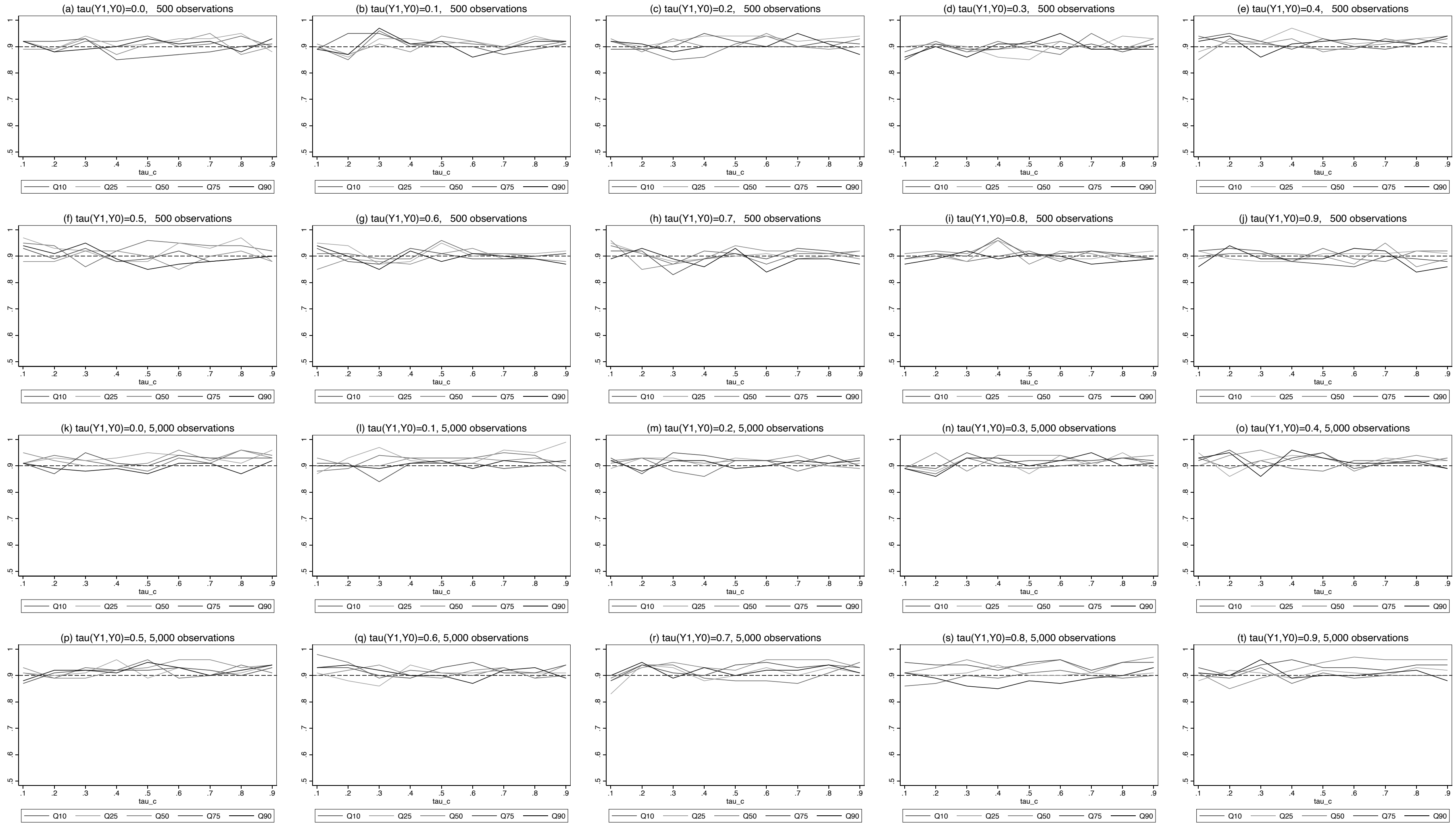


Figure B6: Monte Carlo simulation (100 replications): 90% coverage rate for selected quantiles of GQTE



Appendix C – Proofs

Proof of Theorem 1: Using Lemma 1 of Firpo (2007), the distribution functions $F_{Y(d)}(y)$, $d = \{0,1\}$, can be expressed in terms of weighted averages: $F_{Y(d)}(y) = \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{Y_i \leq y\}$, with $w_{i1} = \frac{D_i}{Pr(D_i=1|X_i)}$ and $w_{i0} = \frac{1-D_i}{1-Pr(D_i=1|X_i)}$. The order statistics $Z_{(1)d} \leq \ldots \leq Z_{(N_{\text{sp}}^d)d}$ are the values of the quantile functions derived from continuous and monotonically increasing distribution functions $F_{Y(d)}(y)$,

$$Z_{(i)d} \quad = \quad F^{-1}_{Y(d)}[\theta_d] = \inf\{y : \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{Y_i \leq y\} \geq \theta_d\}, \tag{18}$$

with $(i-1)/N_{\text{sp}}^d < \theta_d \leq i/N_{\text{sp}}^d$, $i \in \{1,\ldots,N_{\text{sp}}^d\}$. Define the $(N_{\text{sp}}^d \times 1)$-vectors $\mathbf{Z}_{\text{sp},d} = (Z_{1d},\ldots,Z_{N_{\text{sp}}^d d})'$. Consider the $(N_{\text{sp}} \times k)$-covariate matrix $\mathbf{X}_{\text{sp}}$, and the corresponding $((N_{\text{sp}}/2) \times k)$-covariate matrix $\mathbf{X}_{\text{sp},0}$ for members of the control group. The predicted outcomes are given by $\widehat{\mathbf{Y}}_{\text{sp}} = \mathbf{X}_{\text{sp}}(\mathbf{X}'_{\text{sp},0}\mathbf{X}_{\text{sp},0})^{-1}\mathbf{X}'_{\text{sp},0}\mathbf{Y}_{\text{sp},0}$. The elements of the $(N_{\text{sp}}\times 1)$-vector $\widehat{\mathbf{Y}}_{\text{sp}} = (\widehat{Y}_1,\ldots,\widehat{Y}_{N_{\text{sp}}})'$ are used to obtain the order statistics $\widehat{Z}_{(i)d} = \inf\{\widehat{y} : \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{\widehat{Y}_i \leq \widehat{y}\} \geq \theta_d\}$ and to define the $(N_{\text{sp}}^d \times 1)$-vectors $\widehat{\mathbf{Z}}_{\text{sp},d} = (\widehat{Z}_{1d},\ldots,\widehat{Z}_{N_{\text{sp}}^d d})'$.

Let $\Delta^Z_{\text{sp},p} = \mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1} - \mathbf{Z}_{\text{sp},0}$. Under Assumption 2, the distribution of treatment effects may be written as

$$F_\Delta(\delta) = F_{\Delta^Z}(\delta^Z), \tag{19}$$

where

$$\Delta^Z \quad = \quad \sum_{p\in\mathcal{P}_{\text{sp}}} \Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)]\Delta^Z_{\text{sp},p}. \tag{20}$$

After controlling for covariates, permutations that do not satisfy the condition $\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})$ occur with a probability of zero under Assumption 2. Therefore,

$$\Delta^Z \quad = \quad \sum_{p\in\mathcal{S}_{\text{sp}|X}} \Pr[F^Z_{\Delta_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)]\Delta^Z_{\text{sp},p}, \tag{21}$$

where $\mathcal{S}_{\text{sp}|X} = \{p \in \mathcal{P}_{\text{sp}} \mid \tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})\}$. Equation (21) follows from equation (20), Assumptions 2 and 4, and from

$$\begin{aligned}
&\textstyle\sum_{p\in\mathcal{P}_{\text{sp}}} \Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}}] \\
= &\sum_{p\in\mathcal{S}_{\text{sp}|X}} \Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] \\
\times\ &\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] \\
+ &\sum_{p\in\mathcal{S}'_{\text{sp}|X}} \Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] \\
\times\ &\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})],
\end{aligned}$$

where $\mathcal{S}'_{\text{sp}|X} = \{p \in \mathcal{P}_{\text{sp}} \mid \tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\}$. Under Assumption 2, $\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] = 0$ for all $p \in \mathcal{P}_{\text{sp}}$. Using Bayes' law,

$$\begin{aligned}
&\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}}] \\
= \ &\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] \\
\times\ &\Pr[\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})]
\end{aligned}$$

for all $p \in \mathcal{S}_{\text{sp}|X}$. Under Assumption 5,

$$\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] = \frac{\mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\}}{\sum_{\mathcal{P}_{\text{sp}}} \mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\}} \tag{22}$$

for all $p \in \mathcal{P}_{\text{sp}}$. All permutations of $\mathbf{Z}_{\text{sp},1}$ that satisfy $\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})$ have a positive probability of occurrence under Assumptions 2 and 4. Therefore, under Assumptions 2, 4, and 5, the distribution of treatment effects is identified through

$$F_\Delta(\delta) \;=\; F_{\Delta^Z}(\delta^Z), \tag{23}$$

where $\Delta^Z = \frac{1}{n_{p|X}} \sum_{p\in\mathcal{S}_{\text{sp}|X}} \Delta^Z_{\text{sp},p}$ with $n_{p|X} = \sum_{p\in\mathcal{P}_{\text{sp}}} \mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\} = \sum_{p\in\mathcal{S}_{\text{sp}|X}} \mathbf{1}\{\tau(\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\}$. Identification of the Population Quantiles of the Distribution of Treatment Effects follows from

$$q_{\Delta,u} \;=\; q_u(F_\Delta(\delta)) = q_u(F_{\Delta^Z}(\delta^Z)). \tag{24}$$

Q.E.D.

Proof of Theorem 2: Using Lemma 1 of Firpo (2007), the distribution functions $F_{Y(d)}(y)$, $d = \{0,1\}$, can be expressed in terms of weighted averages: $F_{Y(d)}(y) = \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{Y_i \leq y\}$, with $w_{i1} = \frac{D_i}{Pr(D_i=1|X_i)}$ and $w_{i0} = \frac{1-D_i}{1-Pr(D_i=1|X_i)}$. Let

$$i' \in \{\{j_{11},\ldots,j_{1N_{\text{sp}}^{(1-d)}}\},\ldots,\{j_{N_{\text{sp}}^d 1},\ldots,j_{N_{\text{sp}}^d N_{\text{sp}}^{(1-d)}}\}\} = \{1,\ldots,N_{\text{sp}}^1 N_{\text{sp}}^0\},$$

and define the order statistics $Z_{(1)d} \leq \ldots \leq Z_{(N_{\text{sp}}^1 N_{\text{sp}}^0)d}$ as values of the quantile functions derived from continuous and monotonically increasing distribution functions $F_{Y(d)}(y)$,

$$Z_{(i')d} = F_{Y(d)}^{-1}[\theta_d] = \inf\{y : \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{Y_i \leq y\} \geq \theta_d\}, \tag{25}$$

with $(i'-1)/(N_{\text{sp}}^1 N_{\text{sp}}^0) < \theta_d \leq i'/(N_{\text{sp}}^1 N_{\text{sp}}^0)$, $i' \in \{1,\ldots,N_{\text{sp}}^1 N_{\text{sp}}^0\}$.

Define the $(N_{\text{sp}}^1 N_{\text{sp}}^0 \times 1)$-vectors

$$\begin{aligned}\mathbf{V}_{\text{sp},d} &= (Z_{(j_{11})d},\ldots,Z_{(j_{1N_{\text{sp}}^{(1-d)}})d},\ldots,Z_{(j_{N_{\text{sp}}^d 1})d},\ldots,Z_{(j_{N_{\text{sp}}^d N_{\text{sp}}^{(1-d)}})d})' \\ &= (\underbrace{Z_{(1)d},\ldots,Z_{(1)d}}_{N_{\text{sp}}^{(1-d)}},\ldots,\underbrace{Z_{(N_{\text{sp}}^d)d},\ldots,Z_{(N_{\text{sp}}^d)d}}_{N_{\text{sp}}^{(1-d)}})',\end{aligned}$$

where

$$Z_{(i)d} = F_{Y(d)}^{-1}[\theta_d] = \inf\{y : \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{Y_i \leq y\} \geq \theta_d\}, \tag{26}$$

with $(i-1)/N_{\text{sp}}^d < \theta_d \leq i/N_{\text{sp}}^d$, $i \in \{1,\ldots,N_{\text{sp}}^d\}$. Define the $(N_{\text{sp}}^d \times 1)$-vectors $\mathbf{Z}_{\text{sp},d} = (Z_{1d},\ldots,Z_{N_{\text{sp}}^d d})'$. Consider the $(N_{\text{sp}} \times k)$-covariate matrix $\mathbf{X}_{\text{sp}}$, and the corresponding $((N_{\text{sp}}/2) \times k)$-covariate matrix $\mathbf{X}_{\text{sp},0}$ for members of the control group. The predicted outcomes are given by $\widehat{\mathbf{Y}}_{\text{sp}} = \mathbf{X}_{\text{sp}}(\mathbf{X}'_{\text{sp},0}\mathbf{X}_{\text{sp},0})^{-1}\mathbf{X}'_{\text{sp},0}\mathbf{Y}_{\text{sp},0}$. The elements of the $(N_{\text{sp}}\times 1)$-vector $\widehat{\mathbf{Y}}_{\text{sp}} = (\widehat{Y}_1,\ldots,\widehat{Y}_{N_{\text{sp}}})'$ are used to obtain the order statistics $\widehat{Z}_{(i)d} = \inf\{\widehat{y} : \frac{1}{N_{\text{sp}}}\sum_{i=1}^{N_{\text{sp}}} w_{id}\mathbf{1}\{\widehat{Y}_i \leq \widehat{y}\} \geq \theta_d\}$ and to define the $(N_{\text{sp}}^d \times 1)$-vectors $\widehat{\mathbf{Z}}_{\text{sp},d} = (\widehat{Z}_{1d},\ldots,\widehat{Z}_{N_{\text{sp}}^d d})'$.

Consider the permutations $\mathbf{\Omega}_{\text{sp},d}\mathbf{V}_{\text{sp},d} = \mathbf{M}_{\text{sp},d}\mathbf{Z}_{\text{sp},d}$, where $\mathbf{\Omega}_{\text{sp},d}$ are $(N_{\text{sp}}^1 N_{\text{sp}}^0 \times N_{\text{sp}}^1 N_{\text{sp}}^0)$-permutation matrices, and where $\mathbf{M}_{\text{sp},d}$ are $(N_{\text{sp}}^1 N_{\text{sp}}^0 \times N_{\text{sp}}^d)$-transformation matrices that transform the $(N_{\text{sp}}^d \times 1)$-vectors $\mathbf{Z}_{\text{sp},d}$ into the $(N_{\text{sp}}^1 N_{\text{sp}}^0 \times 1)$-vectors

$$\mathbf{M}_{\text{sp},d}\mathbf{Z}_{\text{sp},d} = (\underbrace{Z_{1d},\ldots,Z_{1d}}_{N_{\text{sp}}^{(1-d)}},\ldots,\underbrace{Z_{N_{\text{sp}}^d d},\ldots,Z_{N_{\text{sp}}^d d}}_{N_{\text{sp}}^{(1-d)}})'.$$

Let $\Delta_{\text{sp},p}^Z = \mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1} - \mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}$. Under Assumption 2, the distribution of treatment effects may be written as

$$F_\Delta(\delta) = F_{\Delta^Z}(\delta^Z), \tag{27}$$

where

$$\Delta^Z = \sum_{p\in\mathcal{P}_{\text{sp}}} \Pr[F_{\Delta_{\text{sp},p}^Z}(\delta_{\text{sp},p}^Z) = F_\Delta(\delta)]\Delta_{\text{sp},p}^Z. \tag{28}$$

After controlling for covariates, permutations that do not satisfy the condition $\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})$ occur with a probability of zero under Assumption 2. Therefore,

$$\Delta^Z = \sum_{p\in\mathcal{S}_{\text{sp}|X}} \Pr[F_{\Delta_{\text{sp},p}}^Z(\delta_{\text{sp},p}^Z) = F_\Delta(\delta)]\Delta_{\text{sp},p}^Z, \tag{29}$$

where $\mathcal{S}_{\text{sp}|X} = \{p \in \mathcal{P}_{\text{sp}} \mid \tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\}$. Equation (29) follows from equation (28), Assumptions 2 and 4, and from

$$\begin{aligned}
\textstyle\sum_{p\in\mathcal{P}_{\text{sp}}} & \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}}] \\
= & \sum_{p\in\mathcal{S}_{\text{sp}|X}} \Pr\big[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \\
& \tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\big] \\
\times & \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})] \\
+ & \sum_{p\in\mathcal{S}'_{\text{sp}|X}} \Pr\big[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \\
& \tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\big] \\
\times & \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})],
\end{aligned}$$

where $\mathcal{S}'_{\text{sp}|X} = \{p \in \mathcal{P}_{\text{sp}} \mid \tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1},\mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1},\widehat{\mathbf{Z}}_{\text{sp},1})\}$. Under Assumption 2, $\Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) \neq \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})] = 0$ for all $p \in \mathcal{P}_{\text{sp}}$. Using Bayes' law,

$$\begin{aligned}
& \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}}] \\
= \; & \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\mathbf{Z}_{\text{sp},0}) = \tau_{\text{sp}} \mid \\
& \tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})] \\
\times \; & \Pr[\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})]
\end{aligned}$$

for all $p \in \mathcal{S}_{\text{sp}|X}$. Under Assumption 5,

$$\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] = \frac{\mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\}}{\sum_{\mathcal{P}_{\text{sp}}} \mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\}} \tag{30}$$

for all $p \in \mathcal{P}_{\text{sp}}$. All permutations of $\mathbf{Z}_{\text{sp},1}$ that satisfy $\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})$ have a positive probability of occurrence under Assumptions 2 and 4. Therefore, under Assumptions 2, 4, and 5, the distribution of treatment effects is identified through

$$F_\Delta(\delta) \;\; = \;\; F_{\Delta^Z}(\delta^Z), \tag{31}$$

where $\Delta^Z = \frac{1}{n_{p|X}} \sum_{p \in \mathcal{S}_{\text{sp}|X}} \Delta^Z_{\text{sp},p}$ with $n_{p|X} = \sum_{p \in \mathcal{P}_{\text{sp}}} \mathbf{1}\{\Pr[F_{\Delta^Z_{\text{sp},p}}(\delta^Z_{\text{sp},p}) = F_\Delta(\delta)] > 0\} = \sum_{p \in \mathcal{S}_{\text{sp}|X}} \mathbf{1}\{\tau(\mathbf{M}_{\text{sp},1}\mathbf{\Pi}_{\text{sp},p}\mathbf{Z}_{\text{sp},1}, \mathbf{M}_{\text{sp},0}\widehat{\mathbf{Z}}_{\text{sp},0}) = \tau(\mathbf{Z}_{\text{sp},1}, \widehat{\mathbf{Z}}_{\text{sp},1})\}$. Identification of the Population Quantiles of the Distribution of Treatment Effects follows from

$$q_{\Delta,u} \;\; = \;\; q_u(F_\Delta(\delta)) = q_u(F_{\Delta^Z}(\delta^Z)). \tag{32}$$

Q.E.D.