# Quantile Treatment Effects in Regression Discontinuity Designs with Covariates

**Yu-Chin Hsu**

Institute of Economics

Academia Sinica

&

Department of Finance

National Central University

**Chung-Ming Kuan**†

Department of Finance & CRETA

National Taiwan University

**Giorgio Teng-Yu Lo**

Department of Economics

National Tsing Hua University

† Corresponding author. E-mail: ckuan@ntu.edu.tw. No. 1, Sec. 4, Roosevelt Rd., Taipei, 10617 Taiwan.

## Abstract

Estimating treatment effect heterogeneity conditional on a covariate has become an important research direction for regression discontinuity (RD) designs. In this paper, we go beyond conditional average treatment effect and propose a new method to estimate the conditional quantile treatment effect (CQTE) in RD designs. The proposed CQTE estimator is based on local quantile regression and hence remains local to the cutoff. Moreover, it captures the relations between treatment effects and the covariate directly for each quantile of the outcome variable. The estimated CQTE thus allows us to assess treatment effect heterogeneity with respect to covariates and the outcome variable simultaneously. As such, the proposed procedure renders a more comprehensive picture of treatment effect in RD designs than conventional methods.

**JEL classification:** C13, C21

**Keywords**: Conditional treatment effect, Local quantile regression, Quantile treatment effect, Regression discontinuity design

# 1 Introduction

Program evaluators have long suffered problems such as endogeneity and the incapability of conducting randomized experiments. In recent years, regression discontinuity (RD) design has served as a strategy of choice and become one of the most commonly used frameworks for program evaluation. An RD design arises when the outcome variable exhibits discontinuity at a cutoff value which determines the treatment assignment. In a standard (sharp) RD design, an individual is assigned the treatment when a running variable, or assignment variable, exceeds a given cutoff. The observations with running variables just above and below the cutoff are argued to be more or less similar in terms of other factors or characteristics; see Thistlethwaite and Campbell (1960), Hahn, Todd, and Van der Klaauw (2001), and Lee and Lemieux (2010). As a result, other control variables or instrumental variables needed in other approaches are not necessary for RD designs, and the treatment effect could be estimated by the difference in outcomes between those close to the cutoff value. This stark contrast to other approaches has also allowed for RD designs to gain popularity.

Even though the estimation of RD design only requires the information of the outcome variable, running variable, and cutoff value, other variables are in no way useless. In fact, many studies have emphasized *conditional* treatment effect heterogeneity as a concern and estimate the conditional average treatment effect (CATE) by extending the use of covariates; see, for example, Lehrer, Pohl, and Song (2016), Card and Giuliano (2016), Deshpande (2016), and Pinotti (2017). The existing methods used to estimate average treatment effects (ATEs) conditional on a covariate include the subsample regression method and interaction term method; see Hsu and Shen (2016, 2017).

While one may explore conditional treatment effect heterogeneity by incorporating independent variables, different treatment effects may also be found across the distribution of the dependent variable. Such distributional heterogeneous effects can be analyzed by estimating treatment effects across quantiles of the outcome variable. For instance, Deshpande (2016) discovers the positive effect of removing a cash welfare program on earnings is accentuated at the upper quantiles of earnings using the estimator proposed by Frandsen, Frölich, and Melly (2012). Such asymmetry could only be shown by comparing the effects at different quantiles of the outcome distribution. See also, for example, Beshears (2013), Ito (2015), and Bernal, Carpio, and Klein (2015) for more empirical studies involving quantile treatment effects.

In this paper, we propose a new method to estimate the conditional quantile treatment effect (CQTE) in sharp RD designs. Extending from the local linear estimator of Fan and Gijbels (1992) as well as the conditional treatment effect estimator of Hsu and Shen (2016), the proposed CQTE estimator is based on local quantile regression and hence remains local to the cutoff. Therefore, the method has characteristics closely related to Koenker (1994) and Angrist, Chernozhukov, and Fernández-Val (2006). We suggest adopting a parametric approach in applying quantile regression and a local linear estimator in localizing estimation at the cutoff. Moreover, it captures the relations between treatment effects and the covariate directly for each quantile of the outcome variable. Compared with the subsample regression method, this method incorporates the information of whole sample and avoids arbitrary selection of subsamples. In contrast with the analysis of CATE, the estimated CQTE allows us to assess treatment effect heterogeneity with respect to covariates and the outcome variable simultaneously. As such, the proposed procedure renders a more comprehensive picture of treatment effect in RD designs than conventional methods.

The paper is organized as follows. Section 2 sets up the framework of an RD design with covariates and shows identification of CQTE. Section 3 introduces the proposed estimator of CQTE and testing. Section 4 examines the performance of the proposed estimator in a Monte Carlo simulation study. Section 5 concludes. Proofs and technical assumptions are provided in Appendix.

## 2 Framework and Identification

By selecting observations close to the given cutoff of the running variable, RD designs create *locally randomized experiments*, which can be evaluated under arguably mild assumptions. In this section, we first describe the standard identification of (conditional)

ATEs in RD designs. Secondly, we discuss identification and related assumptions of our proposed method for estimating CATEs.

## 2.1 RD Design and CATE

We consider sharp RD designs in which treatment assignment $T$ is completely determined by a running variable $Z$. A policy intervention encourages an individual to receive treatment if the running variable is larger than or equal to a cutoff value $c$. Let $Y$ denote the outcome of interest, and $T = 1(Z \geq c)$ be the dummy variable indicating treatment of an individual. Let $T = 1$ if treated and $T = 0$ if non-treated. We use $Y(0)$ and $Y(1)$ to denote potential outcomes when $T = 0$ and $T = 1$, respectively. We observe $Z$, $T$ and $Y = T \cdot Y(1) + (1 - T) \cdot Y(0)$.

In a sharp RD design, the ATE at the cutoff is

$$\mathrm{E}[Y(1) - Y(0)|Z = c] = \mathrm{E}[Y(1)|Z = c] - \mathrm{E}[Y(0)|Z = c].$$

Since $\mathrm{E}[Y(0)|Z = c]$ is unobserved, one may impose the continuity assumptions at the cutoff, which could be written as

$$\begin{aligned}\mathrm{E}[Y(0)|Z = c] &= \lim\nolimits_{z \nearrow c} \mathrm{E}[Y(0)|Z = z],\\ \mathrm{E}[Y(1)|Z = c] &= \lim\nolimits_{z \searrow c} \mathrm{E}[Y(1)|Z = z].\end{aligned}$$

It is easy to show that

$$\begin{aligned}\mathrm{E}[Y(0)|Z = c] &= \lim\nolimits_{z \nearrow c} \mathrm{E}[Y(0)|Z = z] = \lim\nolimits_{z \nearrow c} \mathrm{E}[Y|Z = z],\\ \mathrm{E}[Y(1)|Z = c] &= \lim\nolimits_{z \searrow c} \mathrm{E}[Y(0)|Z = z] = \lim\nolimits_{z \searrow c} \mathrm{E}[Y|Z = z],\end{aligned}$$

by the continuity assumptions. Thus, RD designs identify ATE at the cutoff by

$$\mathrm{E}[Y(1) - Y(0)|Z = c] = \lim_{z \searrow c} \mathrm{E}[Y|Z = z] - \lim_{z \nearrow c} \mathrm{E}[Y|Z = z],$$

which is the discontinuity gap of the outcome around the cutoff.

As covariates/control variables are not prerequisite for estimation in an RD analysis, traditionally the major purpose of incorporating additional covariates is to verify whether the underlying assumptions of an RD design is fulfilled. That no other factors other than the treatment status affect the outcome of observations with running variables close to the cutoff point is one of the assumptions in the estimation and implies any other variable should not demonstrate a discontinuity at cutoff. Obtaining the information of additional covariates would enable one to check if the covariates are continuous/balanced around the cutoff as assumed. See, for instance, Imbens and Lemieux (2008), Lee and Lemieux (2010) and Choi and Lee (2016) for theoretical discussions and DiNardo and Lee (2004) and Ludwig and Miller (2007) for empirical applications.

Additional to the ATE of all observations close to the cutoff point, researchers may also be curious whether there exist heterogeneous CATEs across different values of an additional covariate. For instance, the effect of legal status on the crime rate of immigrants is likely to be heterogeneous across different types of crime; see Pinotti (2017). The legal cannabis access may have heterogeneous effects on student's course grades with respect to different gender and age; see Marie and Zölitz (2017). Jepsen, Mueser, and Troske (2016) also find that whether the General Educational Development (GED) certification has significant impacts on earnings would depend on gender. The impact of California's electricity rebate program on energy conservation has also been shown to differ when the temperature, air conditioner saturation, and income level change; see Ito (2015).

In the literature, the subsample regression method and interaction term method are applied to CATE conditional on covariate values. The subsample regression method, to our knowledge, is the more frequently used of the two. The method specifies subsamples and implements nonparametric estimation, which would avoid the observations far away from the cutoff value from affecting the estimated effect at the cutoff, for each subgroup. The effects can then be analyzed separately and straightforwardly for each subgroup. Recent empirical work include Buser (2015), Holden (2016), Deshpande (2016), Card and Giuliano (2016), Jepsen et al. (2016), Beach and Jones (2017), Asher and Novosad (2017), Dague, DeLeire, and Leininger (2017), Hoekstra, Puller, and West (2017), Meng (2017), and Pinotti (2017). However, since the whole sample size needs to be cut by at least half for the estimation of each subsample, the method could often lead to reduction in statistical power, especially when the original sample size is not large; see Hsu and Shen (2016) for detailed discussions. Alternatively, the interaction term method is implemented in some empirical studies. This method applies linear regression with interaction terms between the treatment indicator and covariate and allows one to directly assess whether there exist varying ATEs across different values of the covariate by estimating the coefficient of the interaction term. See, for instance, Beland (2015), Ito (2015), and Beach and Jones (2017) for empirical applications. Although the interaction term can be easily included and interpreted in regressions, the parametric estimation results might be affected by observations far from the cutoff, which would distort the local effect estimate of an RD design and would be prone to yield bias under model misspecification.

## 2.2 CQTE and Identification

As treatment effects are likely to vary across quantiles of the outcome distribution, such differences would be masked by an ATE or CATE. Ito (2015) finds that electricity rebate program affects the change in consumption more at lower quantiles of log consumption than upper quantiles. Analogously, studies evaluating programs such as remedial courses and scholarships could also reflect inherent dissimilarities between different levels of outcome. To capture the distributional heterogeneity, we develop a parametric approach instead of a fully nonparametric approach to identifying CQTE.

To preclude treatment effects from conditioning on infinity, let $X$ be a set of covariates with compact support $\mathcal{X} \subset R^{d_x}$. Let $Q_\tau(Y|X=x, Z=z)$ denote the conditional $\tau$-th quantile of $Y$ conditional on the covariate $X$ and running variable $Z$:

$$Q_\tau(Y|X=x, Z=z) = \inf\{y : F(y|X=x, Z=z) \geq \tau\}, \quad \tau \in (0,1),$$

where $F(y|X=x, Z=z)$ is the distribution function of $Y$ conditional on $X=x$ and $Z=z$. Thus, $Q_\tau(Y(0)|X=x, Z=c)$ and $Q_\tau(Y(1)|X=x, Z=c)$ are the conditional $\tau$-th quantile functions of the potentials at cutoff. The $\tau$-th quantile treatment effect conditional on the covariate $X=x$ is thus

$$\text{CQTE}_\tau(x) = Q_\tau(Y(1)|X=x, Z=c) - Q_\tau(Y(0)|X=x, Z=c).$$

Also let $Y(0) = Q_\tau(Y(0)|X=x, Z=z) + \epsilon_{\tau,0}$ and $Y(1) = Q_\tau(Y(1)|X=x, Z=z) + \epsilon_{\tau,1}$.

We impose the following conditions for identification of CQTE.

**Assumption 2.1** *(Identification) For some $\delta > 0$ and $\tau \in (0,1)$, assume that*

*(i) the conditional quantile functions have the parametric forms:*

$$Q_\tau(Y(0)|X = x, Z = c) = x\beta_0(\tau), \quad Q_\tau(Y(1)|X = x, Z = c) = x\beta_1(\tau), \tag{2.1}$$

*where $\beta_0(\tau)$ and $\beta_1(\tau)$ are interior points of $\mathcal{B}$, a compact subset of $R^{d_x}$;*

*(ii) $J_0(\tau) = E[f_{\epsilon_{\tau,0}}(0|X, Z = c)X'X|Z = c]$ and $J_1(\tau) = E[f_{\epsilon_{\tau,1}}(0|X, Z = c)X'X|Z = c]$ are positive definite.*

*(iii) the probability density function (pdf) of $Z$, $f_z(z)$, is bounded away from zero and bounded above in $z$ on $\mathcal{N}_{\delta,z}(c)$;*

*(iv) the pdf of $X$ and $Z$, $f_{xz}(x, z)$, is uniformly continuous in $x \in \mathcal{X}$ and $z \in \mathcal{N}_{\delta,z}(c)$;*

*(v) the conditional densities $f_{Y_0}(y|X = x, Z = z)$ and $f_{Y_1}(y|X = x, Z = z)$ exist and are bounded and uniformly continuous in $y$, $x \in \mathcal{X}$ and $z \in \mathcal{N}_{\delta,z}(c)$;*

Assumption 2.1(i) assumes that the postulated conditional quantile functions are correctly specified. We could weaken the condition such that the specification is correct almost surely in $X$ without changing all the results. This condition implies that $\beta_t(\tau)$ will be the minimizer of $E[\rho_\tau(Y(t) - Xb)|Z = c]$ over $b$; this result is standard in quantile regression. Assumption 2.1(ii) is a rank condition and implies that $\beta_t(\tau)$ will be the unique minimizer. Assumption 2.1(iii)-(v) are needed to identify $E[\rho_\tau(Y(t) - Xb)|Z = c]$ because $Y(t)$ is not observed all the time. Note that Assumption 2.1(iii) implies that the conditional distribution of $X$ given $Z$ is continuous at the cutoff point which means that $X$ are predetermined variable that is not affected by the treatment status. This is needed for the causal interpretation of the CQTE. Assumption 2.1(iii) also implies that $X$ can includes only continuous variables. This is imposed for notational simplicity and note that our results could be easily extended to cases where $X$ includes discrete variables. The identification result is summarized in the following theorem.

**Theorem 2.1** *Suppose that Assumption 2.1 holds. Then* *(a) for $t \in \{0,1\}$, $\beta_t(\tau)$ is the unique minimizer to the following problem:*

$$\beta_t(\tau) = \arg\min_{b\in\mathcal{B}} E[\rho_\tau(Y(t) - Xb)|Z = c]; \tag{2.2}$$

*(b) $E[\rho_\tau(Y(0) - Xb)|Z = c]$ and $E[\rho_\tau(Y(1) - Xb)|Z = c]$ are identified by*

$$E[\rho_\tau(Y(0) - Xb)|Z = c] = \lim_{z\searrow c} E[\rho_\tau(Y - Xb)|Z = z],$$
$$E[\rho_\tau(Y(1) - Xb)|Z = c] = \lim_{z\nearrow c} E[\rho_\tau(Y - Xb)|Z = z];$$

*(c) in addition,*

$$\lim_{z\searrow c}\sup_{b\in\mathcal{B}} \big|E[\rho_\tau(Y(0) - Xb)|Z = c] - E[\rho_\tau(Y - Xb)|Z = z]\big| = 0,$$
$$\lim_{z\nearrow c}\sup_{b\in\mathcal{B}} \big|E[\rho_\tau(Y(1) - Xb)|Z = c] - E[\rho_\tau(Y - Xb)|Z = z]\big| = 0.$$

The second part of Theorem 2.1 gives the identification of $E[\rho_\tau(Y(t)-Xb)|Z = c]$ for $t = 0, 1$. This is in general obtained by the continuity of $E[\rho_\tau(Y(t)-Xb)|Z = c]$ which is standard in the RD literature. Note that Assumption 2.1 is sufficient for this result to hold. The last part of Theorem 2.1 shows that the identification is uniform in that sense that $E[\rho_\tau(Y - Xb)|Z = z]$ will uniformly converge to $E[\rho_\tau(Y(0) - Xb)|Z = c]$ from below. This is needed so that $\beta_0(\tau; z) = \arg\min_{b\in\mathcal{B}} E[\rho_\tau(Y - Xb)|Z = z]$ will converge to $\beta_0(\tau; c) = \beta_0(\tau)$ when $z$ is approaching $c$ from below. This result is new in the literature.

# 3 CQTE: Estimation, Asymptotics and Inference

In this section, we introduce a CQTE estimator for RD designs. The proposed procedure would allow one to detect heterogeneity with respect to the value of the dependent variable and independent variable.

## 3.1 Estimation of $\beta_1(\tau)$, $\beta_0(\tau)$ and $\text{CQTE}_\tau(x)$

Following the identification result in Theorem 2.1, we propose the following two-step estimator for $\beta_t(\tau)$ for $t = 0, 1$. In the first step, we use local linear estimators to estimate $E[\rho_\tau(Y(t) - Xb)|Z = c]$ for each $b \in \mathcal{B}$. In the second step, we obtain $\beta_t(\tau)$ as the minimizer over estimated $E[\rho_\tau(Y(t) - Xb)|Z = c]$. To be specific, let $\hat{\rho}_{\tau,t}(b)$ be the estimator for $E[\rho_\tau(Y(t)-Xb)|Z=c]$ that solves the following minimization problems:

$$(\hat{\rho}_{\tau,0}(b), \hat{a}^{(\kappa)}_{n,+}(b)) = \arg\min_{\rho,a} \sum_{Z_i<c} K\left(\frac{Z_i-c}{h}\right)\Big[\rho_\tau(Y_i-X_ib)-\rho-a(Z_i-c)\Big]^2,$$

$$(\hat{\rho}_{\tau,1}(b), \hat{a}^{(\kappa)}_{n,-}(b)) = \arg\min_{\rho,a} \sum_{Z_i\geq c} K\left(\frac{Z_i-c}{h}\right)\Big[\rho_\tau(Y_i-X_ib)-\rho-a(Z_i-c)\Big]^2. \quad (3.1)$$

where $K(\cdot)$ is a symmetric kernel function, $h$ is a bandwidth. In the second step, $\hat{\beta}_1(\tau)$ and $\hat{\beta}_0(\tau)$ are obtained by solving the following minimization problems:

$$\hat{\beta}_0(\tau) = \arg\min_{b\in\mathcal{B}} \hat{\rho}_{\tau,0}(b), \quad \hat{\beta}_1(\tau) = \arg\min_{b\in\mathcal{B}} \hat{\rho}_{\tau,1}(b). \quad (3.2)$$

Note that $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$ are equivalent to weighted quantile regression estimators. To see this, first let $S^+_{n,l} = \sum_i^n \mathbf{1}(Z_i \geq c)K(\frac{Z_i-c}{h})(Z_i-c)^l$, $S^-_{n,l} = \sum_i^n \mathbf{1}(Z_i < c)K(\frac{Z_i-c}{h})(Z_i-c)^l$ for $l=0,1,2$, and let

$$w^+_{ni} = \frac{\mathbf{1}(Z_i\geq c)K(\frac{Z_i-c}{h})[S^+_{n,2}-S^+_{n,1}(Z_i-c)]}{S^+_{n,0}S^+_{n,2}-S^+_{n,1}S^+_{n,1}},$$

$$w^-_{ni} = \frac{\mathbf{1}(Z_i< c)K(\frac{Z_i-c}{h})[S^-_{n,2}-S^-_{n,1}(Z_i-c)]}{S^-_{n,0}S^-_{n,2}-S^-_{n,1}S^-_{n,1}},$$

Following Fan and Gijbels (1992), we have

$$\hat{\rho}_{\tau,0}(b) = \sum_{i=1}^n w^-_{ni}\cdot\rho_\tau(Y_i-X_ib), \quad \hat{\rho}_{\tau,1}(b) = \sum_{i=1}^n w^+_{ni}\cdot\rho_\tau(Y_i-X_ib). \quad (3.3)$$

Therefore, $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$ are equivalent to the following weighted quantile regression estimators:

$$\hat{\beta}_0(\tau) = \arg\min_{b\in\mathcal{B}} \sum_{i=1}^n w^-_{ni}\cdot\rho_\tau(Y_i-X_ib), \quad \hat{\beta}_1(\tau) = \arg\min_{b\in\mathcal{B}} \sum_{i=1}^n w^+_{ni}\cdot\rho_\tau(Y_i-X_ib). \quad (3.4)$$

Once we obtain $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$, $\text{CQTE}_\tau(x)$ can in turn be estimated by

$$\widehat{\text{CQTE}}_\tau(x) = x(\hat{\beta}_1(\tau)-\hat{\beta}_0(\tau)). \quad (3.5)$$

### 3.2 Asymptotics of $\hat{\beta}_0(\tau)$, $\hat{\beta}_1(\tau)$ and $\widehat{\textbf{CQTE}}_\tau(x)$

We derive the asymptotics of the proposed estimators in this section. Let $\mu_{\tau,d}(b,z) = E[\rho_\tau(Y(t)-Xb)|Z=z]$.

**Assumption 3.1** *Assume that there exists $\delta > 0$ such that* *(i)* $f_z(z)$ *is twice continuously differentiable in* $z$ *on* $\mathcal{N}_{\delta,z}(c)$*;*

*(ii)* *for* $t = 0, 1,$ *and* $b \in \mathcal{B}$*,* $\mu_d(b; z)$ *is twice continuously differentiable in* $z$ *on* $\mathcal{N}_{\delta,z}(c)$*;*

*(iii)* *for* $t = 0, 1,$ $|\partial^2 \mu_d(b; z)/\partial z \partial z|$ *is uniformly bounded on* $b \in \mathcal{B}$ *and* $z \in \mathcal{N}_{\delta,z}(c)$*;*

**Assumption 3.2** *Assume that*

*(i)* *The function* $K(\cdot)$ *is a non-negative symmetric bounded kernel with a compact support.*

*(ii)* $\int K(u)du = 1$*,*

*(iii)* $h \to 0$*,* $nh \to \infty$ *and* $nh^5 \to 0$ *as* $n \to \infty$*.*

Assumption 3.2(i) is a standard assumption on the kernel function. The triangular kernel ($K(u) = (1 - |u|) \cdot 1(|u| \leq 1)$), which is the most frequently used kernel function in RD estimation and testing, satisfies the conditions stated in the assumption. Assumption 3.2(ii) is the standard undersmoothing condition for local linear estimation. It helps eliminate the nuisance bias term and obtain centered asymptotic normality results of the local linear estimators.

**Assumption 3.3** *Assume that* $f_{\epsilon_{\tau,0}}(0|X = x, Z = c) = f_{\epsilon_{\tau,0}}(0|Z = c)$ *and* $f_{\epsilon_{\tau,1}}(0|X = x, Z = c) = f_{\epsilon_{\tau,1}}(0|Z = c)$ *for all* $x \in \mathcal{X}$*.*

For $j = 0, 1, 2$, let $\vartheta_j = \int_0^\infty u^j K(u)du$. Define $C_k = (\int_0^\infty (\vartheta_2 - u\vartheta_1)^2 K^2(u)du)/\big((\vartheta_2\vartheta_0 - \vartheta_1^2)^2\big)$ and $\Sigma_{xx} = E[X'X|Z = c]$.

**Theorem 3.1** *Suppose that Assumptions 2.1, 3.1 and 3.2 hold. Then*

$$\sqrt{nh}(\hat{\beta}_0(\tau) - \beta_0(\tau)) \xrightarrow{d} \mathcal{N}\big(0, \mathcal{V}_0\big), \quad \sqrt{nh}(\hat{\beta}_1(\tau) - \beta_1(\tau)) \xrightarrow{d} \mathcal{N}\big(0, \mathcal{V}_1\big),$$

*where*

$$\begin{aligned} \mathcal{V}_0 &= \frac{C_k \cdot \tau \cdot (1 - \tau)}{f_z(c)} J_0^{-1}(\tau) \Sigma_{xx} J_0^{-1}(\tau), \qquad (3.6) \\ \mathcal{V}_1 &= \frac{C_k \cdot \tau \cdot (1 - \tau)}{f_z(c)} J_1^{-1}(\tau) \Sigma_{xx} J_1^{-1}(\tau). \end{aligned}$$

*In addition, suppose that Assumption 3.3 also holds, then* $\mathcal{V}_1$ *and* $\mathcal{V}_0$ *reduce to*

$$\mathcal{V}_0 = \frac{C_k \cdot \tau \cdot (1-\tau)}{f_z(c)} \frac{1}{f^2_{\epsilon_{\tau,0}}(0|Z=c)} \Sigma_{xx}^{-1}, \tag{3.7}$$
$$\mathcal{V}_1 = \frac{C_k \cdot \tau \cdot (1-\tau)}{f_z(c)} \frac{1}{f^2_{\epsilon_{\tau,1}}(0|Z=c)} \Sigma_{xx}^{-1}.$$

Theorem 3.1 shows the asymptotic normalities of the proposed estimators $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$. The analytic forms of asymptotic variance-covariance matrices are also given and this result allows us to provide sample analog estimators that are consistent. Theorem 3.1 is similar to standard qunatile regression method except that we consider the local linear regression estimation at the boundary as in the standard RD estimations.

The following corollary regarding the asymptotics of CQTE estimator follows directly from Theorem 3.1.

**Corollary 3.1** *Suppose that Assumptions 2.1, 3.1 and 3.2 hold. Then for* $x \in \mathcal{X}$*,*

$$\sqrt{nh}(\widehat{CQTE}_\tau(x) - CQTE_\tau(x)) \xrightarrow{d} \mathcal{N}\big(0, x'(\mathcal{V}_1 + \mathcal{V}_0)x\big).$$

### 3.3 Inference

To make inference, it is essential to provide consistent estimators for $\mathcal{V}_0$ and $\mathcal{V}_1$. Let $h_1$ be another bandwidth. Similar to Koenker (1994) and Angrist et al. (2006), we estimate $J_0(\tau)$ and $J_1(\tau)$ by

$$\widehat{J}_0(\tau) = \sum_{i=1}^n w_{ni}^+ \cdot 1(|Y_i - X_i'\hat{\beta}_0(\tau)| \le h_1) \cdot X_i X_i' \tag{3.8}$$
$$\widehat{J}_1(\tau) = \sum_{i=1}^n w_{ni}^- \cdot 1(|Y_i - X_i'\hat{\beta}_1(\tau)| \le h_1) \cdot X_i X_i'.$$

For $l = 0, 1, 2$, define

$$S_{n,l} = \sum_{i=1}^n K(\frac{Z_i - c}{h})(Z_i - c)^l,$$
$$w_{ni} = K(\frac{Z_i - c}{h})[S_{n,2} - S_{n,1}(Z_i - c)]/(S_{0,2}S_{n,2} - S_{n,1}^2).$$

Then we estimate $\Sigma_{xx}$ by

$$\widehat{\Sigma}_{xx} = \sum_{i=1}^n w_{ni} \cdot X_i X_i'. \tag{3.9}$$

Then $f_z(c)$ is estimated by a kernel estimator

$$\hat{f}_z(c) = \frac{1}{n}\sum_{i=1}^{n} K\Big(\frac{Z_i - c}{h}\Big). \tag{3.10}$$

Therefore, we can estimate $\mathcal{V}_0$ and $\mathcal{V}_1$ by

$$\widehat{\mathcal{V}}_0 = \frac{C_k \cdot \tau \cdot (1-\tau)}{\hat{f}_z(c)} \widehat{J}_0^{-1}(\tau)\widehat{\Sigma}_{xx}\widehat{J}_0^{-1}(\tau), \tag{3.11}$$
$$\widehat{\mathcal{V}}_1 = \frac{C_k \cdot \tau \cdot (1-\tau)}{\hat{f}_z(c)} \widehat{J}_1^{-1}(\tau)\widehat{\Sigma}_{xx}\widehat{J}_1^{-1}(\tau).$$

In addition, if Assumption 3.3 also holds, then we can first estimate $f_{\epsilon_{\tau,1}}(0|Z = c)$ and $f_{\epsilon_{\tau,0}}(0|Z = c)$ by

$$\hat{f}_{\epsilon_{\tau,0}}(0|Z = c) = \sum_{i=1}^{n} w_{ni}^{-} \cdot 1(|Y_i - X_i'\hat{\beta}_0(\tau)| \le h_1), \tag{3.12}$$
$$\hat{f}_{\epsilon_{\tau,1}}(0|Z = c) = \sum_{i=1}^{n} w_{ni}^{+} \cdot 1(|Y_i - X_i'\hat{\beta}_1(\tau)| \le h_1),$$

and then estimate $\mathcal{V}_0$ and $\mathcal{V}_1$ by

$$\widehat{\mathcal{V}}_0 = \frac{C_k \cdot \tau \cdot (1-\tau)}{\hat{f}_z(c)} \frac{1}{\hat{f}^2_{\epsilon_{\tau,0}}(0|Z = c)} \widehat{\Sigma}_{xx}^{-1}, \tag{3.13}$$
$$\widehat{\mathcal{V}}_1 = \frac{C_k \cdot \tau \cdot (1-\tau)}{\hat{f}_z(c)} \frac{1}{\hat{f}^2_{\epsilon_{\tau,1}}(0|Z = c)} \widehat{\Sigma}_{xx}^{-1},$$

**Assumption 3.4** *Assume that* $h_1 \to 0$ *and* $nh_1^2 \to \infty$.

**Theorem 3.2** *Suppose that Assumptions 2.1, 3.1, 3.2 and 3.4 hold. Then, for* $t = 0, 1$, $\widehat{\mathcal{V}}_t$ *given in (3.11) is consistent for* $\mathcal{V}_t$ *in 3.6. In addition, suppose that Assumption 3.3 also holds, then for* $t = 0, 1$, $\widehat{\mathcal{V}}_t$ *given in (3.13) is also consistent for* $\mathcal{V}_t$.

Theorem 3.2 shows the consistency of the proposed $\widehat{\mathcal{V}}_t$ estimators. The proof is similar to A.1.4. of Angrist et al. (2006).

# 4 Simulation

To examine the plausibility of the proposed method for estimating CQTEs, results of our Monte Carlo simulation study are shown in this section. First of all, we consider sharp RD designs with an outcome variable $Y$, a running variable $Z$, and a covariate $X$. Let

$X$, $Z$, and $\epsilon$ all have $Uni[0,1]$ and are mutually independent. Let treatment variable $T = \mathbf{1}(Z \geq c)$ with cutoff $c = 0.5$.

Four data generating processes (DGPs) are considered in this study. In the first two DGPs, we assume homoskedasticity and let

$$Y = \beta_0 + \beta_1 X + \beta_2 T + \beta_3 TX + \beta_4 Z + \beta_5 ZX + \epsilon. \tag{4.1}$$

Then by construction, the conditional quantile functions of potential outcomes when $Z = c$ are

$$Q_\tau(Y(0)|X = x, Z = c) = (\beta_0 + 0.5\beta_4 + \tau) + (\beta_1 + 0.5\beta_5)x,$$
$$Q_\tau(Y(1)|X = x, Z = c) = (\beta_0 + \beta_2 + 0.5\beta_4 + \tau) + (\beta_1 + \beta_3 + 0.5\beta_5)x.$$

According to (4.1), two scenarios are described as follows.

**DGP 1:** Let $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1, \beta_4 = 0, \beta_5 = 0$.

**DGP 2:** Let $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1, \beta_4 = 2, \beta_5 = 2$.

In DGP 1, the outcome is a linear combination of a constant term, covariate, and error term with different slopes on two sides of the cutoff. In DGP 2, we add the effect of the running variable and its interaction terms with the covariate to the process. The design mimics a situation where how the running variable influences quantile functions depends on values of covariate. Take Jepsen et al. (2016) as an example, it is likely that the decisive GED score affects earnings differently with respect to different values of covariates, such as age and gender.

In DGP 3 and DGP 4, we assume $\epsilon$ varies across values of $X$ and let

$$Y = \beta_0 + \beta_1 X + \beta_2 T + \beta_3 TX + \beta_4 Z + \beta_5 ZX + (1 + X)\epsilon. \tag{4.2}$$

Then by construction, the conditional quantile functions of potential outcomes when $Z = c$ are

$$Q_\tau(Y(0)|X = x, Z = c) = (\beta_0 + 0.5\beta_4 + \tau) + (\beta_1 + 0.5\beta_5 + \tau)x,$$
$$Q_\tau(Y(1)|X = x, Z = c) = (\beta_0 + \beta_2 + 0.5\beta_4 + \tau) + (\beta_1 + \beta_3 + 0.5\beta_5 + \tau)x.$$

According to (4.2), two scenarios are described as follows.

**DGP 3:** Let $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1, \beta_4 = 0, \beta_5 = 0$.

**DGP 4:** Let $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \beta_3 = 1, \beta_4 = 2, \beta_5 = 2$.

The independent variables share the same effect on the outcome with DGP 1 and DGP 2, respectively, except that the error terms now interact with covariate $X$. With the error term being a function of $X$, the robustness to heteroskedasticity can be examined.

We consider a sample size of 5000 with 500 replications. The CQTE estimators are implemented for the 0.25-, 0.5-, 0.75-th quantiles using triangular kernels. All the bandwidths are selected according to the bandwidth algorithm proposed by McCrary (2008). For each quantile, we report the standard error, normalized bias, and mean squared error. Normalized bias is calculated by $\frac{1}{500}\sum_{j=1}^{500}\sqrt{5000h_j}(\hat{\beta}_j(\tau) - \beta(\tau))$ where $h_j$ is the selected bandwidth, $\hat{\beta}_j(\tau)$ is the estimated coefficient for the $j$-th replication and $\beta(\tau)$ is the true coefficient. The mean squared error (MSE) is calculated by $\frac{1}{500}\sum_{j=1}^{500}(\hat{\beta}_j(\tau) - \beta(\tau))^2$.

As shown in Table 1, the CQTE estimation method produces accurate estimates under DGP 1. For all quantiles, the standard errors, normalized biases, and MSEs are all rather small. When the sample size increases, the normalized biases and absolute biases generally decrease. Occasional exceptions reflect slight increase or no decrease in bias when we increase the sample size, but of small magnitude. As expected, all standard errors and MSEs shrink when the sample size increases. Under DGP 2, the results presented in Table 2 remain close to the true parameters. When we add interaction terms of $X$ and $Z$ here, the standard errors and MSEs do not change much, but normalized biases of most of the estimates become slightly larger compared to DGP 1. Similarly, increasing the sample size would reduce the standard errors and MSEs.

As for the heteroskedastic scenarios, the estimates indicate similar patterns. In Table 3, even when the process involves heteroskedasticity, the CQTE estimates under DGP 3 are close to the true parameters. The standard errors, normalized biases, and MSEs are slightly larger than DGP 1, yet still rather small. As we include both heteroskedasticity and interactions terms of $X$ and $Z$, the estimates shown in Table 4 do not differ much from the previous results, indicating heteroskedasticity would hardly affect the estimation of CQTE. All the standard errors and MSEs decrease when we enlarge the sample size under both DGP 3 and DGP 4.

## 5 Conclusion

In this paper, we consider CQTEs for evaluating the treatment effect heterogeneity in sharp RD design. We propose to fit linear quantile regression models for the potential outcomes locally to the cutoff point of the running variable. We derive the asymptotics of the proposed estimators and provide valid inference. The advantage of our method compared to fully nonparametric CQTE estimator is that it is not subject to curse of dimensionality, but it might be subject to model misspecification. For future studies, it is of interest to extend our results to fuzzy RD design and to extend our pointwise result to process results so we can construct confidence band over a continuum of quantile indexes.

# APPENDIX

**Proof of Theorem 2.1:** Part (a) is standard in the quantile regression, so we omit the proof. We show Part (c) first. We show the $t = 1$ case and the arguments for $t = 0$ case are similar. Note that Assumption 2.1 (v) implies when $c \leq z \leq c + \delta$, for all $b \in \mathcal{B}$,

$$|E[\rho_\tau(Y(1) - Xb)|X = x, Z = z] - E[\rho_\tau(Y(1) - Xb)|X = x, Z = c]| \leq M \cdot |z - c|,$$

for some $M < \infty$ because the conditional pdf is uniformly continuous in $z$. This implies that

$$\begin{aligned}
&|E[\rho_\tau(Y(1) - Xb)|Z = z] - E[\rho_\tau(Y(1) - Xb)|Z = c]| \\
=&\Big|\int_{\mathcal{X}} E[\rho_\tau(Y(1) - Xb)|X = \tilde{x}, Z = z]\frac{f_{xz}(\tilde{x}, z)}{f_z(z)}d\tilde{x} \\
&- \int_{\mathcal{X}} E[\rho_\tau(Y(1) - Xb)|X = \tilde{x}, Z = c]\frac{f_{xz}(\tilde{x}, c)}{f_z(c)}d\tilde{x}\Big| \\
\leq& \int_{\mathcal{X}} \big|E[\rho_\tau(Y(1) - Xb)|X = \tilde{x}, Z = z] - E[\rho_\tau(Y(1) - Xb)|X = \tilde{x}, Z = z]\big|\frac{f_{xz}(\tilde{x}, z)}{f_z(z)}d\tilde{x} \\
&+ \int_{\mathcal{X}} E[\rho_\tau(Y(1) - Xb)|X = \tilde{x}, Z = c]\Big|\frac{f_{xz}(\tilde{x}, z)}{f_z(z)} - \frac{f_{xz}(\tilde{x}, c)}{f_z(c)}\Big|d\tilde{x} \\
\leq& M \cdot |Z - c| + M \cdot |Z - c| = 2 \cdot M \cdot |Z - c|,
\end{aligned}$$

where the first equality holds by the definition of conditional expectation and the first inequality holds by triangular inequalities. The second inequality holds by the previous equation and the fact that the $E[\rho_\tau(Y(1) - Xb)|X = x, Z = c]$ is bounded and the conditional density of $f_{x|z}(x|z)$ is uniformly continuous from Assumption 2.1(iii) and (iv). Note that this result holds uniformly over $b$. Then by the fact that $Y = Y(1)$ when $z \geq c$, Part (c) follows. Last, Part (b) is implied by Part (c) and this completes our proof. $\square$

**Proof of Theorem 3.1:** We first show the consistency of $\hat{\beta}_1(\tau)$, i.e., $\hat{\beta}_1(\tau) \xrightarrow{p} \beta_1(\tau)$. To show this, it is essential to establish that

$$\sum_{b \in \mathcal{B}} \big|\hat{\rho}_{\tau,1}(b) - E[\rho_\tau(Y(1) - Xb)|Z = c]\big| \xrightarrow{p} 0.$$

Note that $\{\rho_\tau(Y_i - X_i b) : b \in \mathcal{B},\ 1 \leq i \leq n, n \geq 1\}$ is manageable with respect to a constant function $M$ for some $0 < M < \infty$. The definition of manageability is given in Pollard (1990). Then it follows that $\{w_{ni}^+ \cdot \rho_\tau(Y_i - X_i b) : b \in \mathcal{B},\ 1 \leq i \leq n, n \geq 1\}$

is also manageable. Then by the same argument as in Hsu and Shen (2016), we have $\sqrt{nh}(\hat{\rho}_{\tau,1}(b) - E[\rho_\tau(Y(1) - Xb)|Z = c]$ weakly converge to a tight Gaussian process and this is sufficient to show the the uniform consistency of the $\hat{\rho}_{\tau,1}(b)$ for $E[\rho_\tau(Y(1) - Xb)|Z = c]$. Therefore, this is sufficient to show that the consistency of $\hat{\beta}_1(\tau)$. Once the consistency is established, the rest of the arguments are standard in the quantile regression literature such as Theorem 3 of Angrist et al. (2006) and we omit the details. □

**Proof of Corollary 3.1:** The result follows Theorem 3.1 and the fact that $\sqrt{nh}(\hat{\beta}_0(\tau) - \beta_0(\tau))$ and $\sqrt{nh}(\hat{\beta}_1(\tau) - \beta_1(\tau))$ are independent. They are independent because they are estimated from different subsamples. □

**Proof of Theorem 3.2:** To show this, we show the estimators for $t = 0, 1$, $\widehat{J}_t(\tau)$, $\widehat{\Sigma}_{xx}$, $\hat{f}_z(c)$ and $\hat{f}_{\epsilon_{\tau,t}}(0|Z = c)$ are all consistent. Note that the consistency of $\widehat{\Sigma}_{xx}$ and $\hat{f}_z(c)$ follows from the arguments for local linear regression estimation and kernel estimation, respectively. Note that the estimation effect of $\hat{\beta}_t(\tau)$ will be asymptotically negligible because it converges at a faster rate than $\hat{f}_{\epsilon_{\tau,t}}(0|Z = c)$. Under Assumption 3.3, the consistency of $\hat{f}_{\epsilon_{\tau,t}}(0|Z = c)$ follows directly from Fan, Yao, and Tong (1996). The consistency of $\widehat{J}_t(\tau)$ follows a similar proof for $\hat{f}_{\epsilon_{\tau,t}}(0|Z = c)$, so we omit the details. □

## References


ANGRIST, J., V. CHERNOZHUKOV, AND I. FERNÁNDEZ-VAL (2006): "Quantile regression under misspecification, with an application to the US wage structure," *Econometrica*, 74, 539–563.

ASHER, S. AND P. NOVOSAD (2017): "Politics and Local Economic Growth: Evidence from India," *American Economic Journal: Applied Economics*, 9, 229–273.

BEACH, B. AND D. B. JONES (2017): "Gridlock: Ethnic diversity in government and the provision of public goods," *American Economic Journal: Economic Policy*, 9, 112–136.

BELAND, L.-P. (2015): "Political Parties and Labor-Market Outcomes: Evidence from US States," *American Economic Journal: Applied Economics*, 7, 198–220.

BERNAL, N., M. A. CARPIO, AND T. J. KLEIN (2015): "The effects of access to health insurance for informally employed individuals in Peru," .

BESHEARS, J. (2013): "The performance of corporate alliances: Evidence from oil and gas drilling in the Gulf of Mexico," *Journal of Financial Economics*, 110, 324–346.

BUSER, T. (2015): "The effect of income on religiousness," *American Economic Journal: Applied Economics*, 7, 178–195.

CARD, D. AND L. GIULIANO (2016): "Can Tracking Raise the Test Scores of High-Ability Minority Students?" *The American Economic Review*, 106, 2783–2816.

CHOI, J.-Y. AND M.-J. LEE (2016): "Regression discontinuity: review with extensions," *Statistical Papers*, 1–30.

DAGUE, L., T. DELEIRE, AND L. LEININGER (2017): "The Effect of Public Insurance Coverage for Childless Adults on Labor Supply," *American Economic Journal: Economic Policy*, 9, 124–154.

DESHPANDE, M. (2016): "Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls," *American Economic Review*, 106, 3300–3330.

DiNardo, J. and D. S. Lee (2004): "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001," *The Quarterly Journal of Economics*, 119, 1383–1441.

Fan, J. and I. Gijbels (1992): "Variable bandwidth and local linear regression smoothers," *The Annals of Statistics*, 2008–2036.

Fan, J., Q. Yao, and H. Tong (1996): "Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems," *Biometrika*, 83, 189–206.

Frandsen, B. R., M. Frölich, and B. Melly (2012): "Quantile treatment effects in the regression discontinuity design," *Journal of Econometrics*, 168, 382–395.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

Hoekstra, M., S. L. Puller, and J. West (2017): "Cash for Corollas: When Stimulus Reduces Spending," *American Economic Journal: Applied Economics*, 9, 1–35.

Holden, K. L. (2016): "Buy the Book? Evidence on the Effect of Textbook Funding on School-Level Achievement," *American Economic Journal: Applied Economics*, 8, 100–127.

Hsu, Y.-C. and S. Shen (2016): "Testing for treatment effect heterogeneity in regression discontinuity design," Tech. rep., Institute of Economics, Academia Sinica, Taipei, Taiwan.

——— (2017): "Monotonicity Test for Local Average Treatment Effects Under Regression Discontinuity," Tech. rep., Institute of Economics, Academia Sinica, Taipei, Taiwan.

Imbens, G. W. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, 142, 615–635.

Ito, K. (2015): "Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program," *American Economic Journal: Economic Policy*, 7, 209–37.

JEPSEN, C., P. MUESER, AND K. TROSKE (2016): "Labor Market Returns to the GED Using Regression Discontinuity Analysis," *Journal of Political Economy*, 124, 621–649.

KOENKER, R. (1994): *Confidence Intervals for Regression Quantiles*, Heidelberg: Physica-Verlag HD, 349–359.

LEE, D. S. AND T. LEMIEUX (2010): "Regression discontinuity designs in economics," *Journal of economic literature*, 48, 281–355.

LEHRER, S. F., R. V. POHL, AND K. SONG (2016): "Targeting Policies: Multiple Testing and Distributional Treatment Effects," Working Paper 22950, National Bureau of Economic Research.

LUDWIG, J. AND D. L. MILLER (2007): "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics*, 122, 159–208.

MARIE, O. AND U. ZÖLITZ (2017): ""High" Achievers? Cannabis Access and Academic Performance," *The Review of Economic Studies*, 84, 1210–1237.

MCCRARY, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

MENG, K. C. (2017): "Using a Free Permit Rule to Forecast the Marginal Abatement Cost of Proposed Climate Policy," *American Economic Review*, 107, 748–784.

PINOTTI, P. (2017): "Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime," *American Economic Review*, 107, 138–168.

POLLARD, D. (1990): "Empirical Processes: Theory and Applications," *NSF-CBMS Regional Conference Series in Probability and Statistics*, 2, i–86.

THISTLETHWAITE, D. L. AND D. T. CAMPBELL (1960): "Regression-discontinuity analysis: An alternative to the ex post facto experiment," *Journal of Educational psychology*, 51, 309–317.

Table 1: CQTE Estimates under DGP 1

| | | $\overbrace{\beta_0+0.5\beta_4+\tau}$ | | $\overbrace{\beta_1+0.5\beta_5}$ | | $\overbrace{\beta_0+\beta_2+0.5\beta_4+\tau}$ | | $\overbrace{\beta_1+\beta_3+0.5\beta_5}$ | |
|---|---|---|---|---|---|---|---|---|---|
| **Quantile** | Sample Size | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 |
| $\tau$= **0.25** | Effect | 1.247 | 1.249 | 1.009 | 1.011 | 2.255 | 2.252 | 1.994 | 1.997 |
| | s.e. | (0.094) | (0.062) | (0.154) | (0.109) | (0.102) | (0.064) | (0.172) | (0.110) |
| | N. bias | -0.069 | -0.044 | 0.162 | 0.324 | 0.092 | 0.064 | -0.118 | -0.091 |
| | MSE | 0.009 | 0.004 | 0.024 | 0.012 | 0.010 | 0.004 | 0.030 | 0.012 |
| $\tau$= **0.5** | Effect | 1.493 | 1.495 | 1.015 | 1.011 | 2.496 | 2.502 | 2.002 | 1.996 |
| | s.e. | (0.116) | (0.074) | (0.202) | (0.127) | (0.123) | (0.074) | (0.207) | (0.129) |
| | N. bias | -0.150 | -0.141 | 0.293 | 0.342 | -0.062 | 0.044 | 0.036 | -0.120 |
| | MSE | 0.014 | 0.005 | 0.041 | 0.016 | 0.015 | 0.005 | 0.043 | 0.017 |
| $\tau$= **0.75** | Effect | 1.742 | 1.746 | 1.003 | 1.001 | 2.738 | 2.749 | 2.004 | 1.993 |
| | s.e. | (0.102) | (0.067) | (0.177) | (0.117) | (0.105) | (0.064) | (0.178) | (0.111) |
| | N. bias | -0.167 | -0.120 | 0.077 | 0.045 | -0.223 | -0.035 | 0.066 | -0.205 |
| | MSE | 0.011 | 0.004 | 0.031 | 0.014 | 0.011 | 0.004 | 0.032 | 0.012 |

Table 2: CQTE Estimates under DGP 2

| | | $\widehat{\beta_0 + 0.5\beta_4 + \tau}$ | | $\widehat{\beta_1 + 0.5\beta_5}$ | | $\widehat{\beta_0 + \beta_2 + 0.5\beta_4 + \tau}$ | | $\widehat{\beta_1 + \beta_3 + 0.5\beta_5}$ | |
|---|---|---|---|---|---|---|---|---|---|
| **Quantile** | Sample Size | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 |
| $\tau$= **0.25** | Effect | 2.265 | 2.257 | 1.985 | 1.993 | 3.266 | 3.257 | 3.040 | 3.048 |
| | s.e. | (0.105) | (0.069) | (0.182) | (0.121) | (0.082) | (0.055) | (0.137) | (0.088) |
| | N. bias | 0.293 | 0.204 | -0.299 | -0.206 | 0.309 | 0.238 | 0.783 | 1.434 |
| | MSE | 0.011 | 0.005 | 0.033 | 0.015 | 0.007 | 0.003 | 0.020 | 0.010 |
| $\tau$= **0.5** | Effect | 2.506 | 2.505 | 1.981 | 1.984 | 3.503 | 3.501 | 3.008 | 3.007 |
| | s.e. | (0.110) | (0.073) | (0.180) | (0.124) | (0.110) | (0.078) | (0.184) | (0.128) |
| | N. bias | 0.131 | 0.154 | -0.401 | -0.527 | 0.042 | 0.015 | 0.174 | 0.232 |
| | MSE | 0.012 | 0.005 | 0.033 | 0.016 | 0.012 | 0.006 | 0.034 | 0.016 |
| $\tau$= **0.75** | Effect | 2.736 | 2.744 | 1.960 | 1.950 | 3.751 | 3.751 | 2.994 | 2.990 |
| | s.e. | (0.080) | (0.050) | (0.135) | (0.086) | (0.103) | (0.073) | (0.186) | (0.124) |
| | N. bias | -0.259 | -0.195 | -0.801 | -1.496 | 0.016 | 0.020 | -0.106 | -0.311 |
| | MSE | 0.007 | 0.003 | 0.020 | 0.010 | 0.011 | 0.005 | 0.035 | 0.016 |

Table 3: CQTE Estimates under DGP 3

| | | $\widehat{\beta_0 + 0.5\beta_4 + \tau}$ | | $\widehat{\beta_1 + 0.5\beta_5 + \tau}$ | | $\widehat{\beta_0 + \beta_2 + 0.5\beta_4 + \tau}$ | | $\widehat{\beta_1 + \beta_3 + 0.5\beta_5 + \tau}$ | |
|---|---|---|---|---|---|---|---|---|---|
| **Quantile** | Sample Size | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 |
| $\tau$= **0.25** | Effect | 1.254 | 1.258 | 1.260 | 1.245 | 2.268 | 2.264 | 2.228 | 2.238 |
| | s.e. | (0.112) | (0.080) | (0.251) | (0.168) | (0.124) | (0.079) | (0.262) | (0.171) |
| | N. bias | 0.052 | 0.239 | 0.233 | -0.138 | 0.335 | 0.414 | -0.414 | -0.356 |
| | MSE | 0.012 | 0.006 | 0.063 | 0.028 | 0.016 | 0.006 | 0.069 | 0.029 |
| $\tau$= **0.5** | Effect | 1.496 | 1.509 | 1.507 | 1.485 | 2.511 | 2.511 | 2.463 | 2.478 |
| | s.e. | (0.136) | (0.090) | (0.278) | (0.193) | (0.143) | (0.089) | (0.306) | (0.186) |
| | N. bias | -0.083 | 0.255 | 0.153 | -0.419 | 0.221 | 0.313 | -0.703 | -0.628 |
| | MSE | 0.018 | 0.008 | 0.077 | 0.037 | 0.020 | 0.008 | 0.095 | 0.035 |
| $\tau$= **0.75** | Effect | 1.747 | 1.751 | 1.737 | 1.742 | 2.752 | 2.756 | 2.713 | 2.731 |
| | s.e. | (0.122) | (0.081) | (0.254) | (0.174) | (0.118) | (0.074) | (0.262) | (0.153) |
| | N. bias | -0.063 | 0.038 | -0.236 | -0.209 | 0.044 | 0.165 | -0.693 | -0.539 |
| | MSE | 0.015 | 0.007 | 0.065 | 0.030 | 0.014 | 0.005 | 0.070 | 0.024 |

Table 4: CQTE Estimates under DGP 4

| | | $\overbrace{\beta_0 + 0.5\beta_4 + \tau}$ | | $\overbrace{\beta_1 + 0.5\beta_5 + \tau}$ | | $\overbrace{\beta_0 + \beta_2 + 0.5\beta_4 + \tau}$ | | $\overbrace{\beta_1 + \beta_3 + 0.5\beta_5 + \tau}$ | |
|---|---|---|---|---|---|---|---|---|---|
| **Quantile** | Sample Size | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 | N=2000 | N=5000 |
| $\tau$= **0.25** | Effect | 2.266 | 2.255 | 2.230 | 2.254 | 3.268 | 3.259 | 3.263 | 3.258 |
| | s.e. | (0.135) | (0.081) | (0.265) | (0.170) | (0.106) | (0.070) | (0.221) | (0.141) |
| | N. bias | 0.282 | 0.126 | -0.351 | 0.138 | 0.335 | 0.287 | 0.288 | 0.239 |
| | MSE | 0.019 | 0.007 | 0.070 | 0.029 | 0.012 | 0.005 | 0.049 | 0.020 |
| $\tau$= **0.5** | Effect | 2.502 | 2.505 | 2.489 | 2.498 | 3.505 | 3.493 | 3.482 | 3.505 |
| | s.e. | (0.144) | (0.092) | (0.290) | (0.192) | (0.145) | (0.090) | (0.309) | (0.183) |
| | N. bias | 0.008 | 0.127 | -0.158 | -0.035 | 0.087 | -0.191 | -0.299 | 0.127 |
| | MSE | 0.021 | 0.009 | 0.084 | 0.037 | 0.021 | 0.008 | 0.095 | 0.034 |
| $\tau$= **0.75** | Effect | 2.735 | 2.739 | 2.742 | 2.739 | 3.743 | 3.742 | 3.743 | 3.753 |
| | s.e. | (0.110) | (0.072) | (0.222) | (0.144) | (0.133) | (0.080) | (0.270) | (0.166) |
| | N. bias | -0.308 | -0.348 | -0.149 | -0.324 | -0.145 | -0.216 | -0.084 | 0.099 |
| | MSE | 0.012 | 0.005 | 0.049 | 0.021 | 0.018 | 0.006 | 0.073 | 0.027 |