Quantile Regression with Group-level Treatments

Songnian Chen*
Zhejiang University

December 2022

Abstract

To study the distributional effects of group level treatments, Angrist and Lang (2004) applied quantile regression with group level regressors, and Chetverikov et al. (2016) proposed a grouped instrumental variables quantile regression estimator, a quantile extension of the Hausman and Taylor's (1981) instrumental variables estimator for panel data. However, the analyses of distributional effects of group level treatments in Angrist and Lang (2004) and Chetverikov et al. (2016) are incomplete and their models are quite restrictive, and they only allow for heterogenous distributional effects of group-level treatments that corresponds to individual-level unobserved characteristics, but not group-level unobserved characteristics. In other words, Angrist and Lang (2004) and Chetverikov et al. (2016) allow for within group hetergeneous distributional treatment effects, but not between group heterogeneous distributional treatment effects. In this article, we provide a comprehensive analysis by proposing a quantile regression model that allows for heterogenous distributional effects of group level treatments associated with both individual level and group level unobserved characteristics, corresponding to within-group and between-group distributional effects. We propose two step quantile regression and instrumental variables quantile regression estimators, depending on whether the group level treatments are correlated with the group level unobserved characteristics. Large sample properties are presented and simulation results indicate our estimators perform well in finite samples.

JEL Classification: C21, C23, C26

Key Words: Quantile Regression, Group-level treatments, Endogeneity

1 Introduction

Fixed effects panel data models have played an important role in controlling for time-invariant unobservables. However, the commonly used fixed effect regression estimator fails

*I would like to thank participants at econometric seminar to Georgetown Univerity and econometrics conferences at SHUFE, Xiamen University, and Zhongshan University for helpful comments.

to identify the effects of time-invariant variables that are correlated with time-invariant unobservables. In an influential article, Hausman and Taylor (1981) proposed an instrumental variable method in the context of mean regression to estimate the average effects of the time-invariant variables even in cases without external instruments. To investigate the distributional effects of group level treatments that are potentially correlated with group level unobservables, Angrist and Lang (2004) applied quantile regression with group level regressors in an empirical setting, and Chetverikov et al. (2016) proposed an instrumental variables quantile regression estimator with group level treatments, a quantile extension of the Hausman and Taylor's (1981) instrumental variables estimator for panel data.

However, the analyses of distributional effects of group level treatments in Angrist and Lang (2004) and Chetverikov et al. (2016) are incomplete and their models are quite restrictive, and they only allow for heterogenous distributional effects of group-level treatments that corresponds to individual-level unobserved characteristics, but not group-level unobserved characteristics. In other words, Angrist and Lang (2004) and Chetverikov et al. (2016) only allow for within group heterogenous distributional treatment effects, but not between group heterogeneous distributional treatment effects. In this article, we provide a comprehensive analysis by proposing a quantile regression model that allows for heterogenous distributional effects of group level treatments associated with both individual level and group level unobserved characteristics, allowing for both within group and between heterogenous distributional treatment effects. We propose two step quantile regression and instrumental variables quantile regression estimators, depending on whether the group level treatments are correlated with the group level unobserved characteristics.

Following Chetverikov et al. (2016), we adopt the terminology in the context of group-level treatments, referring panel units as groups and within-group observations as micro-level observations. To highlight the restrictive nature of the framework in Angrist and Lang (2004) and Chetverikov et al. (2016) and to motivate our model and our estimators, we use the example of Boston's Metco program of school integration, which studied how suburban student test scores were affected by the reassignment of participating urban students to suburban schools. Angrist and Lang (2004) and Chetverikov et al. (2016) analyzed the relationship between the presence of urban students in the classroom and the second decile[1] of student test scores by estimating the equation

$$Q_{igjt|x_{gjt}}(0.2) = \alpha_g(0.2) + \beta_j(0.2) + \gamma_t(0.2) + \delta(0.2)m_{gjt} + \lambda(0.2)s_{gjt} + \xi_{gjt}(0.2) \qquad (1.1)$$

where the left-hand side represents the second decile of student test scores within a group, $x_{gjt} = (m_{gjt}, s_{gjt}, \xi_{gjt})$, and a group is a grade $g\times$school $j\times$year t cell. Here m_{gjt} and s_{gjt}

[1]As Metco students' scores mostly concentrated in the lower tail of the score distribution, increasing the fraction of the Metco students shifts the overall score distribution more sharply in the lower tail.

denote the class size and the fraction of Metco students within each $g \times j \times t$ cell, and α_g, β_j, and γ_t represent grade, school, and year effects, respectively. The component ξ_{gjt} represents the unobserved group-level characteristics, which enters the model with an additive separable structure. It is clear that the framework of Angrist and Lang (2004) and Chetverikov et al. (2016) is highly restrictive in the way the group-level unobserved characteristics are allowed to affect the micro-level outcomes. Specifically, for a particular student whose ability placed him at uth percentile within a group based on her individual-level unobserved characteristics, while her affiliated group is placed at τth percentile across all groups based on unobserved group-level unobserved characteristics, their framework (1.1) implies

$$Q_{igjt|x_{gjt},\xi_{gjt}=\tau}(u,\tau) = \alpha_g(u) + \beta_j(u) + \gamma_t(u) + \delta(u)m_{gjt} + \lambda(u)s_{gjt} + \xi(u,\tau), \quad (1.2)$$

which states that the effects of the group-level variables m_{gjt} and s_{gjt} are homogenous across different groups, ruling out general heterogenous distributional effects of group-level policy variables across different groups. In particular, the model of Angrist and Lang (2004) and Chetverikov et al. (2016) can only accommodate general within group heterogeneous distributional effects, but not between heterogeneous distributional effects.

In Section 2, we develop a new quantile regression framework that overcomes the highly restrictive nature of the framework of Angrist and Lang (2004) and Chetverikov et al. (2016). In particular, we allow for general within group and between group heterogeneous distributional effects of group-level treatments, which can vary at different regions of the conditional distribution of the micro-level outcome variable, depending on individual-level and group-level unobserved characteristics; specifically, compared with quantile regression model (1.2) of Angrist and Lang (2004) and Chetverikov et al. (2016), our proposed model corresponds to the following conditional quantile specification:

$$Q_{igjt|x_{gjt},\xi_{gjt}=\tau}(u,\tau) = \alpha_g(u) + \beta_j(u) + \gamma_t(u) + \delta(u,\tau)m_{gjt} + \lambda(u,\tau)s_{gjt} + \xi(u,\tau), \quad (1.3)$$

where, for any individual within group ranking u according to her unobserved individual characteristics, the effects of group variables m_{gjt} and s_{gjt} on the micro-level outcome, $\delta(u,\tau)$ and $\lambda(u,\tau)$, can also vary with τ, which measures the individual's associated group ranking across all groups according to the group-level unobserved characteristics. In particular, $\lambda(u,\tau)$ measures the impact of the school integration program on the individual whose is ranked at u within her group and her affiliated group is ranked at τ among all groups, according to unobserved individual-level and group-level characteristics. We do not impose any parametric structure on $\delta(u,\tau)$ and $\lambda(u,\tau)$. In contrast, Angrist and Lang (2004) and Chetverikov et al. (2016) impose the highly restrictive homogeneity condition that $\delta(u,\tau) = \delta(u)$ and $\lambda(u,\tau) = \lambda(u)$ for all $\tau \in (0,1)$, which implies, for example, that the median (or any other quantiles) student within each group would respond uniformly to group-level treatments, regardless of her group affiliation.

In practice, as the group-level treatments can either be exogenous or endogenous, depending on whether they are correlated with group level unobserved characteristics. Exogenous group treatments are not uncommon, especially in the context of randomized controlled experiments; see, for example, Krueger, 1999, Angrist and Lavy (2009), Angrist and Guryan (2008) and Larsen (2015). On the other hand, endogenous group-level treatments are a common concern in many empirical studies; see Chetverikov et al. for some detailed discussions.

In Section 2, aside from presenting our model, we also discuss its connection to panel data quantile regression in the existing literature. In Section 3, we introduce our quantile regression estimators. Similar to Angrist and Lang (2004) and Chetverikov et al. (2016), our estimators consist of two-steps: (i) in the first step we conduct quantile regression for each group; and (ii) in the second step, we use the estimated group-specific effects as the dependent variable to perform quantile regression for exogenous group-level covariates and instrumental variables quantile regression for endogenous group-level covariates. For the endogenous case, similar to Hausman and Taylor (1981) and Chetverikov et al. (2016), we can also exploit internal instruments when external instruments are difficult to find. In Section 3 we present large sample properties of our estimators. Section 4 contains some Monte Carlo simulation results, which indicate that our estimators perform well in finite samples. Section 5 concludes. Proofs of the main theorems are in the Appendix.

2 The Model

In this section we propose a quantile regression model that focuses on micro-level response to group-level treatments. Our model allows for flexible distributional effects arising from both unobserved individual-level characteristics and unobserved group-level characteristics.

To motivate our model, we first describe the set-up of Chetverikov et al. (2016)

$$Q_{y_{ig}}(u|\tilde{z}_{ig}, x_g, \varepsilon_g) = \tilde{z}_{ig}'\gamma(u) + \tilde{x}_g'\tilde{\beta}\,(u) + \varepsilon_g(u) = z_{ig}'\alpha_g(u), \quad u \in \mathcal{U} \tag{2.1}$$

where y_{ig} is the micro-level response variable of individual i in group g, $Q_{y_{ig}}(u|\tilde{z}_{ig}, \tilde{x}_g, \varepsilon_g)$ is the uth conditional quantile of y_{ig} given $\tilde{z}_{ig}, \tilde{x}_g, \varepsilon_g$, $\tilde{z}_{ig}$ is a $(d_z - 1)$-vector of observable individual-level covariates, $\tilde{x}_g$ is a $(d_x - 1)$-vector of observable group-level covariates[2], $\gamma(u)$ and $\tilde{\beta}\,(u)$ are the respectively conformable vectors of uth quantile regression coefficients, $\varepsilon_g = \{\varepsilon_g(u),\, u \in \mathcal{U}\}$ is a set of unobservable group-level random scalar shifters, and $\mathcal{U}$ is a set of quantile indices of interest. The main parameter of interest is $\tilde{\beta}(u)$.

[2]Here $\tilde{x}_g$ does not contain a constant term since no location normalization is imposed on $\varepsilon_g\,(u)$. In Chetverikov et al. (2016), x_g contains a constant term and a location normalization is imposed by assuming $Ex_g\varepsilon_g\,(u) = 0$ for the exogenous case and $Ew_g\varepsilon_g(u) = 0$ for the endogenous case with instruments w_g.

To better understand their framework, we restate model (2.1) in terms of the corresponding data generating mechanism:

$$y_{ig} = \tilde{z}_{ig}'\gamma(U_{ig}) + \tilde{x}_g'\tilde{\beta}\left(U_{ig}\right) + \varepsilon(U_{ig}, U_g) = z_{ig}'\alpha_g(U_{ig}, U_g) \tag{2.2}$$

where U_{ig}, U_g describe the individual-level and group-level unobserved characteristics. As in a typical quantile regression set-up U_{ig} is interpreted as the ranking of individual i within group g associated with unobserved individual characteristics, whereas U_g describes the ranking of group g based on its group-level unobserved characteristics across all groups, thus both U_{ig} and U_g have uniform marginals on $(0, 1)$, assumed to be independent of each other. In the context of Angrist and Lang (2004), among the observables, $\tilde{x}_g$ could include the class size and the fraction of the Metco students, and $\tilde{z}_{ig}$ represents the individual and family characteristics; among the unobservables, U_g is a measure of the general learning environment including teacher quality, whereas U_{ig} could represent the student's ability. Model (2.2) allows for general interaction between $(\tilde{z}_{ig}, \tilde{x}_g)$ and U_{ig}, but not between $(\tilde{z}_{ig}, \tilde{x}_g)$ and U_g. Based on (2.2), (2.1) can be equivalently expressed as

$$Q_{y_{ig}}(u|\tilde{z}_{ig}, x_g, \varepsilon_g) = \tilde{z}_{ig}'\gamma(u) + \tilde{x}_g'\tilde{\beta}\left(u\right) + \varepsilon(u, U_g) = z_{ig}'\alpha_g(u), \quad u \in \mathcal{U}. \tag{2.3}$$

Notice that (2.3) imposes a highly restrictive homogeneity feature that the quantile coefficients $\tilde{\beta}(u)$ is uniform across different groups; in particular, the group-level unobserved characteristics U_g only affects the micro-level outcome variable y_{ig} through a location shift, and thus the policy effect of the group-level variable $\tilde{x}_g$ is uniform across different groups, regardless of the group-level unobserved heterogeneity. Consequently, while Angrist and Lang (2004) and Chetverikov et al. (2016) allow for within group heterogeneous distributional treatment effects, but not between group heterogeneous distributional treatment effects.

To overcome the above-mentioned drawback associated with the framework in Chetverikov et al. (2016), we propose the following quantile regression model with group-level treatments

$$\begin{aligned} y_{ig} &= \tilde{z}_{ig}'\gamma(U_{ig}) + \tilde{x}_g'\tilde{\beta}\left(U_{ig}, U_g\right) + \varepsilon(U_{ig}, U_g) \\ &= \tilde{z}_{ig}'\gamma(U_{ig}) + x_g'\beta\left(U_{ig}, U_g\right). \end{aligned} \tag{2.4}$$

Define

$$\alpha_{g,1}(U_{ig}, U_g) = x_g'\beta\left(U_{ig}, U_g\right) = \tilde{x}_g'\tilde{\beta}\left(U_{ig}, U_g\right) + \varepsilon(U_{ig}, U_g) \tag{2.5}$$

Then equations (2.4) and (2.5) imply the following individual level and group level conditional quantile regression restrictions:

$$Q_{y_{ig}}(u|\tilde{z}_{ig}, x_g, U_g) = \tilde{z}_{ig}'\gamma(u) + \tilde{x}_g'\tilde{\beta}\left(u, U_g\right) + \varepsilon(u, U_g) \tag{2.6}$$

and

$$Q_{\alpha_{g,1}(u)}(\tau|x_g) = x_g'\beta\left(u,\tau\right) = \tilde{x}_g'\tilde{\beta}\left(u,\tau\right) + \varepsilon_g(u,\tau), \tag{2.7}$$

or equivalently

$$y_{ig}|\left(\tilde{z}_{ig}, x_g, U_{ig} = u, U_g = \tau\right) = \tilde{z}_{ig}'\gamma(u) + \tilde{x}_g'\tilde{\beta}\left(u,\tau\right) + \varepsilon(u,\tau)$$

if x_g is exogenous. When x_g is endogenous and we have appropriate instrument w_g, then (2.4) and (2.5) imply that

$$E\left[1\left\{\alpha_{g,1}(u) < \tilde{x}_g'\tilde{\beta}\left(u,\tau\right) + \varepsilon(u,\tau)\right\}|w_g\right] = \tau. \tag{2.8}$$

Consequently, in contrast to the model proposed by Chetverikov et al. (2016) in which the group-level unobserved characteristics only affect the micro-level outcome variable through a location shift, our model offers much more flexible interaction between observed group level treatments and unobserved micro-level and group-level characteristics. As we do not impose any parametric structure on $\beta\left(u,\tau\right)$, therefore our model allows for very general heterogeneous distributional effects of the group-level treatments $\tilde{x}_g$ on the conditional distribution of the micro-level outcomes[3].

In the discussion of their quantile regression model, Chetverikov et al. (2016, page 813) considered the example of modeling the effects of a group level (state-year cohort) policy $\tilde{x}_g$ on the conditional wage distribution; in their setting, while the policy $\tilde{x}_g$ is allowed to have differential effects on lower wage quantiles (with less unobserved skills) from upper wage quantiles, however, such differential policy effects are only confined to the within group variation, whereas homogeneity is maintained across different groups; for example, for the individuals from different groups who are ranked at uth quantile within their respective groups based on their individual unobserved skills, the policy effects on these individuals are uniform, regardless of their respective between group rankings according to their unobserved group level characteristics. In contrast, our model allows for completely unrestricted differential quantile wage effects that can vary with unobserved individual skills (U_{ig}) as well as unobserved group-level heterogeneity (U_g). In the case of Boston school integration program, let u represent the rankings of a student within her group and τ the ranking of the her group among all groups based on the individual and group level unobserved characteristics, then $\lambda(u,\tau)$ in (1.3), the coefficient of s_{gjt}, the fraction of Metco students for that group, would represent the effect of policy change for this particular student; in particular, we allow the school integration program to have different impacts on students with the same

[3] In a different context, Chesher (2003) and Jun (2009) studied identification and estimation of a parameter $\alpha(\tau_1,\tau_2)$ which measures the return to education for the individual who is ranked at τ_1 and τ_2 in terms of unobserved ability, and unobserved market fortune respectively.

within between group ranking but different group rankings, which is ruled out in the setting of Chetverikov et al. (2016).

In this article we consider both the exogenous and endogenous policy variables $\tilde{x}_g$. As in Chetverikov et al. (2016), we have observations on G groups and N_g individuals within group $g = 1, ..., G$. Thus the data set consists of $\{(z_{ig}, y_{ig}), i = 1, ..., N_g\}$ and $\tilde{x}_g$ for $g = 1, ..., G$; when $\tilde{x}_g$ is considered to be endogenous, we also observe some appropriate instruments w_g. As in Hausman and Taylor (1981) and Chetverikov et al. (2016), internal instruments made up by components of z_{ig} are potentially applicable. This feature is sometimes particularly appealing since external instruments can often be difficult to find.

In a related context, Kato et al. (2012), Galvao and Kato (2016), and Gu and Volgushev (2019), among others, considered the quantile regression model

$$y_{ig} = \tilde{z}_{ig}'\tilde{\gamma}(U_{ig}) + \alpha(U_{ig}).$$

The focus of these papers is the estimation of the quantile coefficients $\gamma(u)$ that corresponds to the micro-level variables. Arellano and Bonhomme (2016) considered a fixed T panel quantile regression that imposes a conditional quantile structure on the relationship between the latent group level unobserved heterogeneity and exogenous covariates. These papers do not consider the estimation of effects of the group-level policy variables $\tilde{x}_g$. Hahn and Meinecke (2005) extended Hausman and Taylor (1981) to a nonlinear setting, but their analysis assumes homogeneous effects of individual level and group-level variables.

3 Estimators

In this section, we develop our estimators for $\beta(u, \tau)$ in (2.7) and (2.8) for $u, \tau \in \mathcal{U} \times \mathcal{T}$, where $\mathcal{U}$ and $\mathcal{T}$ are sets of quantile indices. We consider cases of exogenous and endogenous group-level variables x_g. In each case, our estimator consists of two steps.

In the first step, we run group-by-group smoothed uth quantile regression of y_{ig} on $\tilde{z}_{ig}$ and on a constant to obtain $\hat{\alpha}_{g,1}(u)$, an estimate of $\alpha_{g,1}(u)$. In the second step, we run τth smoothed quantile regression or instrumental variables quantile regression of $\hat{\alpha}_{g,1}(u)$ on x_g for the estimation of $\beta(u, \tau)$. Due to the computational and theoretical advantages[4], we adopt smoothed quantile regression in both steps.

We now describe the details of the first step estimator. Base on the data $\{(z_{ig}, y_{ig}), i = 1, ..., N_g\}$, for each group g and a given quantile index u, first note the objective function of the standard

[4] He et al. (2020) demonstrates that smoothed quantile regression possesses significant computational advantage for very large samples and/or a large number of regressors, together with slightly superior performance. Fernandes et al. (2019) shows that the smoothed quantile regression estimator dominates the standard quantile regression in the AMSE sense. Smoothed quantile regression has also been considered by Horowitz (1998), Galvao and Kato (2016), and Kaplan and Sun (2017), among others.

quantile regression takes the form

$$\hat{R}_g^*(\alpha;u) = \int \rho_u(t)d\hat{F}_g^*(t;\alpha) = \frac{1}{N_g}\sum_{i=1}^{N_g}\rho_\tau(y_{ig} - z_{ig}'\alpha)$$

where $\rho_u(t) = 1\{t > 0\}ut - 1\{t \le 0\}(1-u)t$ and $\hat{F}_g^*(t;\alpha) = \frac{1}{N_g}\sum_{i=1}^{N_g} 1\{y_{ig} - z_{ig}'\alpha \le t\}$. Following Fernandes et al. (2019), we define the smoothed quantile regression estimator for $\alpha_g(u)$ by replacing $\hat{F}_g^*(t;\alpha)$ with a smoothed version, $\hat{F}_g(t;\alpha)$; specifically, we define

$$\hat{\alpha}_g(u) = \arg\min_{\alpha \in R^{d_z}} \hat{R}_g(\alpha, u)$$

where

$$\begin{aligned}
\hat{R}_g(\alpha,u) &= \int \rho_u(t)d\hat{F}_g(t;\alpha) \\
&= \frac{1}{N_g}\sum_{i=1}^{N_g}\int \rho_u(y_{ig} - z_{ig}'\alpha + h_1 v)k_1(v)dv, \\
&= (1-u)\int_{-\infty}^{0}\hat{F}_g(t;\alpha)dt + u\int_0^\infty \left(1 - \hat{F}_g(t;\alpha)\right)dt
\end{aligned}$$

with

$$\hat{F}_g(t;\alpha) = \int_{-\infty}^{t}\hat{f}_g(v;\alpha)dv$$

and

$$\hat{f}_g(v;\alpha) = \frac{1}{N_g h_1}\sum_{i=1}^{N_g} k_1\left(\frac{y_{ig} - z_{ig}'\alpha - v}{h_1}\right).$$

Here $k_1(\cdot)$ and h_1 are the kernel function and bandwidth parameter respectively. Let $\hat{\alpha}_{g,1}(u) = e_1'\hat{\alpha}_g(u)$, which serves as the dependent variable for the second stage estimation, and e_1 is the unit vector with the first element equal to 1.

For the second step estimation, we first consider the case with exogenous x_g. For $\tau \in \mathcal{T}$, we perform smoothed quantile regression of $\hat{\alpha}_{g,1}(u)$ on x_g to obtain $\hat{\beta}^*(u,\tau)$ as our estimator for $\beta(u,\tau)$ for the exogenous case:

$$\hat{\beta}^*(u,\tau) = \arg\min_{b \in R^{d_x}} \hat{R}(b,u,\tau)$$

where

$$
\begin{aligned}
\hat{R}(b,u,\tau) &= \int \rho_\tau(t) d\hat{F}(t;b,u) \\
&= (1-u)\int_{-\infty}^{0} \hat{F}(t;b,u)dt + u\int_{0}^{\infty} \left(1-\hat{F}(t;b,u)\right) dt
\end{aligned}
$$

with

$$
\hat{F}(t;b,u) = \int_{-\infty}^{t} \hat{f}(v;b,u)dv
$$

and

$$
\hat{f}(v;b,u) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2\left(\frac{\hat{\alpha}_{g,1}(u) - x_g'b - v}{h_2}\right).
$$

Here $k_2(\cdot)$ and h_2 are the kernel function and bandwidth parameter for the second stage, respectively.

For the case with endogenous x_g, we propose to estimate $\beta(u,\tau)$ by $\hat{\beta}(u,\tau)$, which solves

$$
\frac{1}{G}\sum_{g=1}^{G}\left(K_2\left(\frac{\hat{\alpha}_{g,1}(u) - x_g'b}{h_2}\right) - \tau\right) w_g = 0, \tag{4.1}
$$

where $K_2(t) = \int_{-\infty}^{t} k_2(s)ds$ and w_g is the vector of appropriate instruments.

Remark: Note that solving (4.1) could be potentially demanding when x_g contains multiple variables. If $\tilde{x}_g$ contains both endogenous and exogenous group variables, then we could adopt a computationally more attractive a procedure. Let $\tilde{x}_g = (\tilde{x}_{g1}, \tilde{x}_{g2})$ where $\tilde{x}_{g1}$ and $\tilde{x}_{g2}$ represent endogenous and exogenous group variables, respectively. Suppose $x_g'b = \tilde{x}_{g1}'b_1 + x_{g2}'b_2$, where $x_{g2} = (1, \tilde{x}_{g2}')'$. Then, given $\hat{\alpha}_{g,1}(u)$, we propose a two-step procedure for the estimation of $\beta(u,\tau)$, in the spirit of Chernozhukov and Hansen (2006, 2008) and Chen (2018). Specifically, for any b_1, $\hat{b}_{2\tau}(b_1)$ minimizes $\hat{R}(b_1, b_2, u, \tau)$ with respect to b_2, where $\hat{R}(b,u,\tau)$ is defined above. In the second step, define $\hat{b}_{1\tau}$ as a solution to the equation

$$
\frac{1}{G}\sum_{g=1}^{G}\left(K_2\left(\frac{\hat{\alpha}_{g,1}(u) - \tilde{x}_{g1}'b_1 - x_{g2}'\hat{b}_{2\tau}(b_1)}{h_2}\right) - \tau\right) w_g = 0
$$

and we further define $\hat{b}_\tau = (\hat{b}_{1\tau}', \hat{b}_{2\tau}'(\hat{b}_{1\tau}))'$ as our estimator for $\beta(u,\tau)$. For the case with exact identification when $\dim(x_g) = \dim(w_g)$, the limiting distributions of the two alternative estimators are the same.

4 Large Sample Properties

In this section we present the large sample properties of our estimators. We consider both the exogenous and endogenous cases. As the endogenous case contains more general results, we first present the large sample results for the group instrumental variables quantile regression estimator.

Let $E_g[\cdot] = E[\cdot|x_g, U_g]$ and $y_{1g}(u) = z_{1g}'\alpha_g(u, U_g) = \tilde{z}_{1g}'\gamma(u) + \alpha_{g,1}(u, U_g)$. Let $F_g(\cdot)$ and $f_g(\cdot)$ denote the conditional distribution and density functions of y_{1g} given (z_{1g}, x_g, U_g) and $f_g(u, \cdot)$ and $f_{\alpha,g}(u, \cdot)$ denote the conditional density functions of $y_{1g}(u)$ and $\alpha_{g,1}(u, U_g)$ given (z_{1g}, x_g, w_g) respectively. Note that the conditional density of $y_{1g}(u)$ at $z_{1g}'\alpha_g(u, \tau)$ given (z_{1g}, x_g, w_g) is the same as that of $\alpha_{g,1}(u, U_g)$ at $\alpha_{g,1}(u, \tau) = x_g'\beta(u, \tau)$. The terms such as c, C, c_f, C_f, c_M, and C_M etc. are generic constants, which can take on different values in different places. $\mathcal{U}$ and $\mathcal{T}$ denote two compact quantile index subsets of $(0, 1)$. We make the following assumptions.

Assumption 1: (i) Observations are independent across groups. (ii) For all $g = 1, ..., G$, the pairs (z_{ig}, y_{ig}) are i.i.d across $i = 1, ..., N_g$, conditional on the group-level variables (x_g, U_g).
Assumption 2: (i) For all $g = 1, ..., G$, and $i = 1, ..., N_g$, random vectors z_{ig} and (x_g, w_g) satisfy $\|z_{ig}\| \leq C_M$, $\|x_g\| \leq C_M$ and $\|w_g\| \leq C_M$. (ii) For all $g = 1, ..., G$, all eigenvalues of $E_g\left[z_{1g}z_{1g}'\right]$ are bounded from below by c_M.
Assumption 3: (i) For all $u \in \mathcal{U}$ and $g = 1, ..., G$, $f_g(\cdot)$ and $f_g(u, \cdot)$ are strictly positive everywhere and s_1th and s_2th order continuously differentiable respectively. In addition, all derivatives are uniformly bounded by C_f, $f_g\left(z_{1g}'\alpha_g(u, U_g)\right) > c_f$ and $f_g(u, z_{1g}'\alpha_g(u, \tau)) > c_f$ for $u \in \mathcal{U}$ and $\tau \in \mathcal{T}$. (ii) For each $\delta > 0$,

$$\varepsilon_\delta = \inf_{g, u \in \mathcal{U}} \inf_{\|\alpha - \alpha_g(u)\| = \delta} E\left[\int_0^{z_{ig}'\alpha} \{F_g(s) - u\}\, ds\right] > 0.$$

Assumption 4: The kernel function k_1 is continuously differentiable up to s_1th order with a bounded support; in addition, it is an s_1th order kernel such that $\int z^j k_1(z)dz = 0$ for $j = 1, ..., s_1 - 1$, and $\int z^{s_1} k_1(z)dz \neq 0$.
Assumption 5: For all $u \in \mathcal{U}$ and $\tau \in \mathcal{T}$, $\beta(u, \tau)$ is an interior point of $\mathcal{B}$, and $\mathcal{B}$ is a compact set.
Assumption 6: As $G \to \infty$, for each $u \in \mathcal{U}$, $\frac{1}{G}\sum_{g=1}^G E\left[\left(1\left\{\alpha_{g,1}(u, U_g) < x_g'b\right\} - \tau\right) w_g\right] \to S(u, \tau, b)$ uniformly over $(\tau, b) \in \mathcal{T} \times \mathcal{B}$; $S(u, \tau, b)$ is continuous in b and $\beta(u, \tau)$ is the unique solution to $S(u, \tau, b) = 0$ for all $u \in \mathcal{U}$ and $\tau \in \mathcal{T}$.
Assumption 7: As $G \to \infty$, $\frac{1}{G}\sum_{g=1}^G E\left[w_g w_g'\right] \to Q_{ww}$, $\frac{1}{G}\sum_{g=1}^G E\left[w_g x_g'\right] \to Q_{wx}$ and

$\frac{1}{G}\sum_{g=1}^{G} E\left[f_{\alpha,g}(u, x_g'b)w_g x_g'\right] \to Q_u(b)$ uniformly over (u,τ,b) for all $u \in \mathcal{U}$, $\tau \in \mathcal{T}$ and $b \in \mathcal{B}$, where Q_{ww}, Q_{wx} and $Q_{u,\tau} = Q_u(\beta(u,\tau))$ are matrices with singular values in absolute value bounded below by c_M and from above by C_M.

Assumption 8: The kernel function k_2 is continuously differentiable up to s_2th order with a bounded support; in addition, it is an s_2th order kernel such that $\int z^j k_2(z)dz = 0$ for $j = 1, ..., s_2 - 1$, and $\int z^{s_2} k_2(z)dz \neq 0$.

Let $\delta_G = N_G^{-1} h_1^{-1/2} \ln^2 N_G + h_1^{s_1}$, where $N_G = \min_g N_g$.

Assumption 9: As $G \to \infty$ and $N_G \to \infty$, $h_1 \to 0$, $h_2 \to 0$, $\delta_G^3 h_2^{-3} = o\left(G^{-1/2}\right)$, $h_2^{s_2} = o\left(G^{-1/2}\right)$, and $N_G^{-1} \times O_p((Gh_2^3)^{-1/2} \ln^2 G + h_2^{-1}) = o_p\left(G^{-1/2}\right)$; in addition, $\lambda_1 \leq \frac{\max_g N_g}{\min_g N_g} \leq \lambda_2$ holds for some positive constants λ_1 and λ_2.

Assumption 1 describes the data generating mechanism, and as in Chetverikov et al. (2016) we assume independence within and across groups. Assumption 2 contains some boundedness condition as well as the full rank condition needed for the first step estimator. Assumption 3(i) contains some boundedness and smoothness conditions, which, together with the conditions on the first stage kernel function and bandwidth parameter in Assumption 4, ensures good asymptotic properties and Bahadur representation of the first step estimator. Assumption 3(ii) is a form of uniform global identification, which is used in the proof of uniform consistency of $\hat{\alpha}_g(u)$ over g and u. Assumptions 1-4 largely follow Fernandes et al. (2019). Assumptions 5 and 6 are needed for global identification and consistency, which is common for moment based estimators. Assumption 7 contains the full rank conditions needed for the second stage group level quantile regression. Assumption 8 imposes some common conditions on the second stage kernel function. Assumption 9 imposes some rather weak conditions on the bandwidth parameters, and the relative growth rates of N_G and G. For example, if we choose $s_1 = s_2 = 2$ and $h_1 = N_g^{-1/3} \ln N_g$ and $h_2 = G^{-1/3} \ln G$, then Assumption 9 is satisfied when $\left(G^{3/4} \ln N_G\right)/N_G \to 0$, which allows the number of observations for each group grows at a slower rate than the number of groups.

The following lemma describes the large sample properties of the first stage quantile regression estimator. The results and their proof largely follow Fernandes et al. (2019).

Lemma 1: Under Assumptions 1-4, $\hat{\alpha}_{g,1}(u)$ is consistent for $\alpha_{g,1}(u)$ and satisfies

$$\hat{\alpha}_{g,1}(u) - \alpha_{g,1}(u)) = \frac{1}{\sqrt{N_g}} \sum_{i=1}^{N_g} \phi_{ig}(u) + r_g(u)$$

with

$$r_g(u) = O_p\left(N_g^{-1} h_1^{-1/2} \ln N_g + h_1^{2s_1}\right)$$

uniformly over $u \in \mathcal{U}$ and $g \in \{1, 2, ..., G\}$, where

$$\phi_{ig}(u) = e_1' J_g^{-1}(u) \left[1\left\{ y_{ig} - z_{ig}' \alpha_g(u) < 0 \right\} - u \right] z_{ig}$$

with $J_g(u) = E_g \left[f_g(z_{1g}' \alpha_g(u)) z_{1g} z_{1g}' \right]$.

The next theorem contains the results on the second stage instrumental variables quantile regression estimator.

Theorem 2: Under Assumptions 1-9, for a given $u \in \mathcal{U}$

$$\begin{aligned}
& \sqrt{G}\left(\hat{\beta}(u,\tau) - \beta(u,\tau)\right) \\
= & \frac{1}{\sqrt{G}} \sum_{g=1}^{G} \left(Q_{u,\tau}' Q_{u,\tau}\right)^{-1} Q_{u,\tau}' \left(1\left\{\alpha_{g,1}(u) < x_g' \beta(u,\tau)\right\} - \tau\right) w_g + o_p(1) \\
= & \frac{1}{\sqrt{G}} \sum_{g=1}^{G} \left(Q_{u,\tau}' Q_{u,\tau}\right)^{-1} Q_{u,\tau}' \left(1\left\{y_{1g}(u) < z_{1g}' \alpha(u,\tau)\right\} - \tau\right) w_g + o_p(1)
\end{aligned}$$

uniformly over $\tau \in \mathcal{T}$, and

$$\sqrt{G}\left(\hat{\beta}(u,\tau) - \beta(u,\tau)\right) \Longrightarrow \mathbb{G}(\cdot), \quad \text{in} \quad l^{\infty}(\mathcal{T})$$

where $\mathbb{G}(\cdot)$ is a zero-mean Gaussian process with uniformly continuous sample paths and covariance function

$$\mathcal{C}_u(\tau_1, \tau_2) = \left(Q_{u,\tau}' Q_{u,\tau}\right)^{-1} Q_{u,\tau}' \Sigma_{\tau\tau'} Q_{u,\tau} \left(Q_{u,\tau}' Q_{u,\tau}\right)^{-1},$$

where $\Sigma(\tau, \tau') = [\min(\tau, \tau') - \tau\tau'] Q_{ww}$.

In order to conduct statistical inferences, we need an estimator for the asymptotic covariance function. Define

$$\widehat{\mathcal{C}}_u(\tau_1, \tau_2) = \left(\hat{Q}_{u,\tau}' \hat{Q}_{u,\tau}\right)^{-1} \hat{Q}_{u,\tau}' \hat{\Sigma}_{\tau\tau'} \hat{Q}_{u,\tau} \left(\hat{Q}_{u,\tau}' \hat{Q}_{u,\tau}\right)^{-1}$$

where

$$\hat{Q}_{u,\tau} = \frac{1}{Gh_2} \sum_{g=1}^{G} k_2 \left(\frac{\hat{\alpha}_{g1}(u) - x_g' \hat{\beta}(u,\tau)}{h_2} \right) w_g x_g'$$

and $\hat{\Sigma}(\tau, \tau') = [\min(\tau, \tau') - \tau\tau'] \hat{Q}_{ww}$ with $\hat{Q}_{ww} = \frac{1}{G} \sum_{g=1}^{G} w_g w_g'$.

Theorem 3: Under Assumptions 1-9, $\left\|\widehat{\mathcal{C}}_u(\tau_1,\tau_2)-\mathcal{C}_u(\tau_1,\tau_2)\right\| = o_p(1)$ uniformly in $\tau_1,\tau_2 \in \mathcal{T}$.

Finally, we present the large sample results for the exogenous case.

Assumption 7': As $G\to\infty$, $\frac{1}{G}\sum_{g=1}^{G} E\left[x_g x_g'\right] \to Q_{xx}$ and $\frac{1}{G}\sum_{g=1}^{G} E\left[f_g(u, z_{1g}'(u,\tau)) w_g x_g'\right] \to Q_{xx,u,\tau}$, where Q_{xx} and $Q_{xx,u,\tau}$ are matrices with singular values in absolute value bounded below by c_M and from above by C_M, uniformly over (u,τ) for all $u\in\mathcal{U}$ and $\tau\in\mathcal{T}$.

Theorem 4: Assume Assumptions 1-4, 8-9 and Assumption 7' hold, then for a given $u\in\mathcal{U}$,

$$\begin{aligned}
&\sqrt{G}\left(\hat{\beta}^*(u,\tau)-\beta(u,\tau)\right)\\
=\ &\frac{1}{\sqrt{G}}\sum_{g=1}^{G} Q_{xx,u,\tau}^{-1}\left(1\left\{\alpha_{g,1}(u) < x_g'\beta(u,\tau)\right\}-\tau\right)x_g + o_p(1)\\
=\ &\frac{1}{\sqrt{G}}\sum_{g=1}^{G} Q_{xx,u,\tau}^{-1}\left(1\left\{y_{1g}(u) < z_{1g}'\alpha(u,\tau)\right\}-\tau\right)x_g + o_p(1)
\end{aligned}$$

uniformly over $\tau\in\mathcal{T}$, and

$$\sqrt{G}\left(\hat{\beta}^*(u,\tau)-\beta(u,\tau)\right) \Longrightarrow \mathbb{G}_x(\cdot), \quad \text{in} \quad l^\infty(\mathcal{T})$$

where $\mathbb{G}_x(\cdot)$ is a zero-mean Gaussian process with uniformly continuous sample paths and covariance function

$$\mathcal{C}_{xx,u}(\tau_1,\tau_2) = Q_{xx,u,\tau}^{-1}\Sigma_{xx\tau\tau'}Q_{u,\tau}^{-1},$$

where $\Sigma_{xx}(\tau,\tau') = \left[\min(\tau,\tau')-\tau\tau'\right]Q_{xx}$.

As the proof Theorem 4 is essentially a special case of that of Theorem 2, the details are omitted. Also note that in the exogenous case, the second stage is also a convex minimization problem; as a result, the parameter space for the group-level coefficients does not need to be compact. In addition, the above results are for a given u, and can be easily extended to a finite number of quantile indices.

5 Simulation Studies

In this section, we examine the finite sample properties of our estimators. In our Monte Carlo experiments, data is generated from the following model

$$\begin{aligned}
y_{ig} &= z_{ig1}\gamma_1(U_{ig}) + z_{ig2}\gamma_2(U_{ig}) + (0.5+x_g)\beta(U_{ig},U_g)\\
x_g &= w_g + \delta\times U_g
\end{aligned}$$

where

$$\gamma_1(U_{ig}) = \gamma_2(U_{ig}) = \Phi^{-1}(U_{ig})$$

and

$$\beta(U_{ig}, U_g) = \Phi^{-1}(U_{ig}) + \Phi^{-1}(U_g),$$

w_g, z_{ig1}, z_{ig2} are each distributed $U(0,4)$; U_{ig} and U_g are both distributed $U(0,1)$ and thus $\Phi^{-1}(U_{ig})$ and $\Phi^{-1}(U_g)$ are standard normal; the random variables w_g, z_{ig1}, z_{ig2}, U_{ig} and U_g are mutually independent; we set δ to 0 for the exogenous case and 0.5 for the endogenous case.

We estimate the quantile regression coefficients for x_g using the two-step quantile regression estimator and the two-step instrumental variable quantile regression estimator proposed in this article for the exogenous case and the endogenous case respectively. We consider the quantile combinations $(u, \tau) = (0.3, 0.5, 0.7) \times (0.3, 0.5, 0.7)$. For the sample sizes, we consider the combinations of the number of groups (G) and group size N_G given by $(G, N_G) = (25, 50)$, $(50, 50)$, $(50, 100)$, $(100, 50)$, $(100, 100)$, $(100, 200)$, $(200, 100)$ and $(200, 200)$. One thousand Monte Carlo replications are used. The results are reported in Tables 1-4.

Table I reports the results for exogenous case, including the biases (Bias), standard errors (SD) and root mean squared errors (RMSE). We follow the implementation of the smoothed quantile regression of Fernandes et al. (2019) as in He et al. (2021) for both the first stage and second stage of the estimation. We adopt their default choice of the bandwidth in the implementation by He et al. (2021), which is quite robust across different designs.

For the combination of ($G = 25$ and $N_G = 50$) with relatively modest G/N, our estimator performs reasonably well with small biases. For $G = 50$, when N_G increases from 50 to 100, the performance of our estimator overall improves slightly, which indicates that when N_G is large enough relative to G, G is the main factor in determining the properties of our estimator, as suggested in our theoretical results. Indeed, compared with the combination of $(G = 50, N_G = 100)$, the combinations $(G = 100, N_G = 50)$ and $(G = 100, N_G = 100)$ produce significantly smaller SDs. This pattern continues when G or N_G or both increases to 200 from 100.

Tables II-IV report the results for the endogenous case. In the first stage, we use the same smoothed quantile regression as in the exogenous case. In the second stage of the instrumental quantile regression, we adopt the standard normal kernel and the bandwidth h_2 is set $c_0 * sd * G^{-1/3}$ where sd denotes the sample standard error of $\hat{\alpha}_{g1}(u) - x_g' b$; we experimented $c_0 = 0.5$, 1 and 1.5, and the results are not sensitive. Our estimator still performs reasonably well, even for the case with $(G = 25, N_G = 50)$, and a similar pattern persists for different combinations of G and N_G, as in the exogenous case. On the other hand, compared with the performance our estimator for the exogenous case, our two-step instrumental variables estimator produces larger biases and standard errors. Both biases and standard errors decrease as the sample size increases.

6 Conclusion

In this article, we have proposed a quantile regression model that allows for heterogenous distributional effects of group level treatments associated with both individual level and group level unobserved characteristics. We propose two step quantile regression and instrumental variables quantile regression estimators, depending on whether the group level treatments are correlated with the group level unobserved characteristics. In this paper we allow for general interactions between group level treatment x_g and the unobservable characteristics (U_{ig}, U_g) at both at the micro and group levels. A more general setup takes the form

$$y_{ig} = \tilde{z}_{ig}'\gamma(U_{ig}, U_g) + \tilde{x}_g'\tilde{\beta}\,(U_{ig}, U_g) + \varepsilon(U_{ig}, U_g)$$

which also allows for general interactions between the observed individual level characteristics z_{ig} and the unobservable characteristics (U_{ig}, U_g). We will consider this more general framework in future research.

Appendix

Proof of Lemma 1: First we establish the uniform consistency of $\hat{\alpha}_g(u)$ across $g = 1, ..., G$ and over $u \in \mathcal{U}$. Let

$$R_g^*(\alpha; u) = E\left[\rho_\tau(y_{ig} - z_{ig}'\alpha)\right]$$

and

$$R_g(\alpha, u) = \int E\left[\rho_u(y_{ig} - z_{ig}'\alpha + h_1 z)\right] k_1(z)dz.$$

Note that

$$\begin{aligned}\hat{R}_g(\alpha, u) - \hat{R}_g(\alpha(u), u) &= \left[(\hat{R}_g(\alpha, u) - \hat{R}_g(\alpha(u), u)) - (R_g(\alpha, u) - R_g(\alpha(u), u))\right] \\ &\quad + \left[(R_g(\alpha, u) - R_g(\alpha(u), u)) - \left(R_g^*(\alpha, u) - R_g^*(\alpha(u), u)\right)\right] \\ &\quad + \left[\left(R_g^*(\alpha, u) - R_g^*(\alpha(u), u)\right)\right]. \qquad \text{(A1)}\end{aligned}$$

For any given δ, let $B_g(\delta, u) = \{\alpha\colon \ \|\alpha - \alpha_g(u)\| \le \delta\}$, namely, $B_g(\delta, u)$ is the ball of radius of δ centered at $\alpha_g(u)$. Then the Knight's identity (Knight, 1998) and Assumption 3(ii) imply that

$$\left|R_g^*(\alpha, u) - R_g^*(\alpha(u), u)\right| \ge \varepsilon_\delta \qquad \text{(A2)}$$

if $\alpha \notin B_g(\delta, u)$. In addition, by Lemma 1 of Fernandes et al. (2019) we have

$$\sup_{\alpha,u,g}\left|(R_g(\alpha, u) - R_g(\alpha(u), u)) - \left(R_g^*(\alpha, u) - R_g^*(\alpha(u), u)\right)\right| = O\left(h_1^{s_1}\right) < \varepsilon_\delta/2 \qquad \text{(A3)}$$

for a large N_G. As $\hat{\alpha}_g(u)$ is a minimizer of $\hat{R}_g(\alpha, u)$, thus

$$\hat{R}_g(\hat{\alpha}_g(u), u) - \hat{R}_g(\alpha(u), u) \le 0. \qquad \text{(A4)}$$

Then from (A1)-(A4) we can deduce

$$\begin{aligned}&\Pr(\max_{g,u}\|\hat{\alpha}_g(u) - \alpha_g(u)\| \ge \delta) \\ \le\ & \Pr\left(\cup_{g=1}^G \sup_{\alpha\in B_g(\delta,u)}\left|(\hat{R}_g(\alpha, u) - \hat{R}_g(\alpha(u), u)) - (R_g(\alpha, u) - R_g(\alpha(u), u))\right| \ge \varepsilon_\delta/2\right) \\ \le\ & G\max_g \Pr\left(\sup_{\alpha\in B_g(\delta,u)}\left|(\hat{R}_g(\alpha, u) - \hat{R}_g(\alpha(u), u)) - (R_g(\alpha, u) - R_g(\alpha(u), u))\right| \ge \varepsilon_\delta/2\right).\end{aligned}$$

Consider the function

$$f_{\alpha,h,u}(y,z) = \int \rho_u(y_{ig} - z_{ig}'\alpha - hs)k_1(s)ds$$

indexed by (α, h, u), and define the class of functions

$$\mathcal{F} = \left\{ f_{\alpha,h,u}(y,z)\colon\ h_1 \in R, \alpha \in R^{d_z} \right\}.$$

To show that $\mathcal{F}$ is Euclidean (Pakes and Pollard, 1989) with an integrable envelop, define the class of functions

$$\mathcal{F}_\rho = \left\{ \rho_u(y_{ig} - z_{ig}'\alpha - h)\colon\ h \in R,\ \alpha \in R^{d_z} \right\}.$$

Lemma 6 in Chetverikov et al (2016) shows that $\mathcal{F}_\rho$ is VC subgraph, and thus Euclidean with an integrable envelop (Lemma 2.12, Pakes and Pollard, 1989), and then from Corollary 21 in Nolan and Pollard (1987), we can deduce that $\mathcal{F}$ is Euclidean with an integrable envelop. By applying the Talagrand's inequality (Bousquet 2002) as in Kato et al. (2012), we obtain

$$\Pr\left(\sup_{\alpha \in B_g(\delta,u)} \left| (\hat{R}_g(\alpha,u) - R_g(\alpha,u) \right| > \varepsilon_\delta/4 \right) \le c_1 \exp\left(-c_2 N_g \varepsilon_\delta^2\right), \tag{A5}$$

and

$$\Pr\left(\left| (\hat{R}_g(\alpha(u), u) - R_g(\alpha(u), u) \right| > \varepsilon_\delta/4 \right) \le c_1 \exp\left(-c_2 N_g \varepsilon_\delta^2\right) \tag{A6}$$

for some positive constant terms c_1 and c_2. Let

$$\Omega_0 = \left\{ \max_g \sup_u |\hat{\alpha}_g(u) - \alpha_g(u)| < \varepsilon_\delta \right\}.$$

From the above result, we obtain

$$\Pr(\Omega_0) > 1 - 2Gc_1 \exp\left(-c_2 N_G \varepsilon_\delta^2\right) > 1 - cN_G^{-3} \tag{A.7}$$

when G and N_G are large enough, where the second inequality follows from Assumption 9. This establishes the uniform consistency of $\hat{\alpha}_g(u)$ over $g = 1, ..., G$ and $u \in \mathcal{U}$.

We now derive the uniform asymptotic representation. Define the events

$$\Omega_1 = \left\{ \sqrt{N_g} \sup_{\alpha,u} \left\| \hat{R}_g^{(1)}(\alpha,u) - R_g^{(1)}(\alpha,u) \right\| \ge \ln N_g \right\}$$

and

$$\Omega_2 = \left\{ \sqrt{N_g h_1} \sup_{\alpha,u} \left\| \hat{R}_g^{(2)}(\alpha,u) - R_g^{(2)}(\alpha,u) \right\| \ge \ln N_g \right\}.$$

where

$$\hat{R}_g^{(1)}(\alpha, u) = \frac{1}{N_g}\sum_{i=1}^{N_g}\left(K_1\left(\frac{y_{ig} - z_{ig}'\alpha}{h_1}\right) - u\right) z_{ig}$$

$$R_g^{(1)}(\alpha, u) = E_g\left[\left(K_1\left(\frac{y_{ig} - z_{ig}'\alpha}{h_1}\right) - u\right) z_{ig}\right]$$

$$\hat{R}_g^{(2)}(\alpha, u) = \frac{1}{N_g}\sum_{i=1}^{N_g}\frac{1}{h_2}k_1\left(\frac{z_{ig}'\alpha - y_{ig}}{h_1}\right) z_{ig}z_{ig}'$$

and

$$R_g^{(2)}(\alpha, u) = E_g\left[\frac{1}{h_1}k_1\left(\frac{y_{ig} - z_{ig}'\alpha}{h_1}\right) z_{ig}z_{ig}'\right].$$

Similar to (A5) and (A6), we can show

$$\Pr(\Omega_1) \leq c_1 \exp(-c_2 \ln^2 N_g) \tag{A8}$$

and

$$\Pr(\Omega_2) \leq c_1 \exp(-c_2 \ln^2 N_g). \tag{A9}$$

From Assumptions 2-4 and 9, we can deduce that $R_g^{(2)}(\alpha, u)$ is uniformly positive definite with its eigenvalues bounded away from zero for a large enough N_G, and thus $\hat{R}_g^{(2)}(\alpha, u)$ is nonsingular on Ω_2 for large enough N_G. Define $\Omega_G = \Omega_0 \cap \Omega_1 \cap \Omega_2$, then by (A7)-(A9) we have $\Pr(\Omega_G) > 1 - cN_G^{-2}$ for large enough G and N_G. Therefore, on the event Ω_G, a Taylor expansion yields

$$\hat{R}_g^{(1)}(\hat{\alpha}_g(u)) = \frac{1}{N_g}\sum_{i=1}^{N_g}\left\{K_1\left(\frac{z_{ig}'\hat{\alpha}_g(u) - y_{ig}}{h_1}\right) - u\right\} z_{ig} = 0,$$

and

$$\hat{\alpha}_g(u) - \alpha_g(u) = -\left[\int_0^1 \hat{R}_g^{(2)}(\alpha_g(u) + t\left[\hat{\alpha}_g(u) - \alpha_g(u)\right]; u)\, dt\right]^{-1} \hat{R}_g^{(2)}(\alpha_g(u); u).$$

Now with arguments similar to those following equation (28) in the proof of Proposition 3 in Fernandes et al (2019), we obtain

$$\hat{\alpha}_{g,1}(u) - \alpha_{g,1}(u)) = \frac{1}{\sqrt{N_g}}\sum_{i=1}^{N_g}\phi_{ig}(u) + r_g(u)$$

with

$$\max_g |r_g(u)| = O_p\left(N_g^{-1} h_1^{-1/2} \ln N_g + h_1^{2s_1}\right)$$

uniformly over $u \in \mathcal{U}$, which establish Lemma 1.

Here we first present some preliminary lemmas, which are used in the proof the main theorems. For Lemmas A1-A5, we implicitly assume that Assumptions 1-9 hold.

Lemma A1: As G and N_G go to infinity,

$$\frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\hat{\alpha}_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right] w_g = S(u, \tau, b) + o_p(1)$$

uniformly over (τ, b) for a given u.

Proof of Lemma A1: Let

$$\hat{S}_G(u, \tau, b) = \frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\hat{\alpha}_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right] w_g.$$

A Taylor expansion yields

$$\hat{S}_G(u, \tau, b) = S_G(u, \tau, b) + s_G(u, \tau, b), \tag{A10}$$

where

$$S_G(u, \tau, b) = \frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\alpha_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right] w_g$$

and

$$s_G(u, \tau, b) = \frac{1}{G}\sum_{g=1}^{G}\left[k_2\left(\frac{\bar{\alpha}_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right] w_g \left(\frac{\hat{\alpha}_{g,1}(u) - \alpha_{g,1}(u)}{h_2}\right)$$

where $\bar{\alpha}_{g,1}(u)$ lies between $\alpha_{g,1}(u)$ and $\hat{\alpha}_{g,1}(u)$. Lemma 1 and Assumptions 8 and 9 imply that

$$\|s_G(u, \tau, b)\| = O_p\left(N_g^{-1} h_1^{-1/2} h_2^{-1} \ln N_g + h_1^{s_1} h_2^{-1}\right) = o_p(1) \tag{A11}$$

uniformly over (u, τ, b).

Now consider $S_G(u, \tau, b)$. Define the class of functions

$$\mathcal{F}_u = \left\{\left(K_2\left(\frac{\alpha_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right) w_g \colon \tau \in \mathcal{T}, b \in R^{d_x}\right\}.$$

By Lemma 22 in Nolan and Pollard (1987), $\mathcal{F}_u$ is Euclidean with an integrable envelop for a fixed u, and thus manageable (Pollard, 1990). Consequently, by the uniform law of large numbers (Pollard, 1990), we obtain

$$S_G(u,\tau,b) = S(u,\tau,b) + o_p(1) \tag{A12}$$

uniformly in (u,τ,b). Combining (A10)-(A12) establishes Lemma A1.

Lemma A2: As G and N_G go to infinity,

$$\hat{Q}_u(b) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2\left(\frac{\hat{\alpha}_{g1}(u) - x_g'b}{h_2}\right) w_g x_g' \to^p Q_u(b)$$

uniformly over $b \in \mathcal{B}$.

Proof of Lemma A2: A Taylor expansion yields

$$\hat{Q}_{G,u}(b) = Q_{G,u}(b) + Q_{G1,u}(b) + Q_{G2,u}(b) \tag{A13}$$

where

$$Q_{G,u}(u,b) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2\left(\frac{\alpha_{g1}(u) - x_g'b}{h_2}\right) w_g x_g',$$

$$Q_{G1,u}(b) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2'\left(\frac{\alpha_{g1}(u) - x_g'b}{h_2}\right) w_g x_g' \left(\frac{\hat{\alpha}_{g1}(u) - \alpha_{g1}(u)}{h_2}\right)$$

and

$$Q_{G2,u}(b) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2''\left(\frac{\bar{\alpha}_{g1}(u) - x_g'b}{h_2}\right) w_g x_g' \left(\frac{\hat{\alpha}_{g1}(u) - \alpha_{g1}(u)}{h_2}\right)^2,$$

where $\bar{\alpha}_{g1}(u)$ lies between $\hat{\alpha}_{g1}(u)$ and $\alpha_{g1}(u)$.

Let

$$Q_{G,u}^{ml}(u,b) = \frac{1}{Gh_2}\sum_{g=1}^{G} k_2\left(\frac{\alpha_{g1}(u) - x_g'b}{h_2}\right) w_{gm} x_{gl}.$$

Define the class of functions $\mathcal{F}_1 = \left\{k_2\left(\frac{\alpha_{g,1}(u) - x_g'b}{h_2}\right) w_{gm}x_{gl}\colon\ h_2 \in R, b \in R^{d_x}\right\}$. By Example 2.10 and Lemma 2.14 of Pakes and Pollard (1989), $\mathcal{F}_1$ is Euclidean with an integrable envelop, thus by Lemma 8.3.1 of Jin and Ying (2004),

$$Q_{G,u}^{ml}(b) = EQ_G^{ml}(u,b) + (Gh_2)^{1/2}\ln G = Q_u^{ml}(b) + o_p(1) \tag{A14}$$

uniformly over $b \in \mathcal{B}$, where the second equality follows from Assumptions 2, 3 and 9. Similarly, we can show

$$\frac{1}{Gh_2}\sum_{g=1}^{G}\left|k_2'\left(\frac{\alpha_{g1}(u)-x_g'b}{h_2}\right)\right| = O_p(1)$$

and

$$\frac{1}{Gh_2}\sum_{g=1}^{G}\left|k_2''\left(\frac{\alpha_{g1}(u)-x_g'b}{h_2}\right)\right| = O_p(1)$$

uniformly over $b \in \mathcal{B}$, which, together with Assumptions 2, 9 and Lemma 1, imply

$$Q_{G1,u}(b) = O_p\left(\delta_G h_2^{-1}\right) = o_p(1) \tag{A15}$$

and

$$Q_{G2,u}(b) = O_p\left(\delta_G^2 h_2^{-3}\right) = o_p(1) \tag{A16}$$

uniformly in b. Combining (A13)-(A16) establishes Lemma A2.

Lemma A3: As G and N_G go to infinity,

$$\begin{aligned} & S_G(u,\tau,\beta(u,\tau)) \\ = & \frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\alpha_{g1}(u)-x_g'\beta(u,\tau)}{h_2}\right)-\tau\right]w_g \\ = & \frac{1}{G}\sum_{g=1}^{G}\left[1\left\{\alpha_{g1}(u)<x_g'\beta(u,\tau)\right\}-\tau\right]w_g + o_p\left(G^{-1/2}\right) \end{aligned}$$

uniformly over $\tau \in \mathcal{T}$.

Proof of Lemma A3: Define

$$s_g(u,b) = w_g K_2\left(\frac{\alpha_{g1}(u)-x_g'b}{h_2}\right) = w_g\int 1\left\{\alpha_{g1}(u)-x_g'b<h_2 v\right\}k_2(v)\,dv$$

and

$$\tilde{s}_g(u,b) = 1\left\{\alpha_{g1}(u)<x_g'b\right\}w_g.$$

Let

$$S_G^{\Delta}(u,b) = \frac{1}{G}\sum_{g=1}^{G}\left(s_g(u,b)-\tilde{s}_g(u,b)\right).$$

Define the class of functions

$$\mathcal{F}_2 = \left\{ s_g(u,b) - \tilde{s}_g(u,b) \colon \; b \in R^{d_x} \right\}.$$

Let $\sigma^2 = \sup_{b,\tau} E\left[s_g(u,b) - \tilde{s}_g(u,b)\right]^2$. With some calculus, it is easy to see that $\sigma^2 = O(h_2)$. As in Lemmas A1 and A2, we can show that $\mathcal{F}_2$ is Euclidean with the constants (A, V) (Pakes and Pollard, 1989), then by Corollary 5.1 of Chernozhukov et al. (2014), we obtain

$$\begin{aligned} & E\left[\sup_b \left| \left(S_G^{\Delta}(u,b) - S^{\Delta}(u,b)\right) - E\left(S_G^{\Delta}(u,b) - S^{\Delta}(u,b)\right)\right|\right] \\ \leq & \sqrt{G\sigma^2 V \log(AC/\sigma^2)} = O\left(G^{-1/2}\right). \end{aligned}$$

In addition, by Assumption 9, we can show that $E\left(S_G^{\Delta}(u,b) - S^{\Delta}(u,b)\right) = O\left(h_2^{s_2}\right) = o(G^{-1/2})$ uniformly in b. By collecting the above results, we establish Lemma A3.

Lemma A4: As G and N_G go to infinity, $\sqrt{G}\hat{S}_{G1}(u,\tau) = o_p(1)$ uniformly over $\tau \in \mathcal{T}$.
Proof of Lemma A4: Recall

$$\hat{S}_{G1}(u,\tau) = \frac{1}{G}\sum_{g=1}^{G} \frac{1}{h_2} k_2\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right) w_g\left(\hat{\alpha}_{g1}(u) - \alpha_{g,1}(u)\right).$$

By Lemma 1, write $\hat{S}_{G1}(u,\tau)$ as

$$\hat{S}_{G1}(u,\tau) = \hat{S}_{G1}^{*}(u,\tau) + R_{G1}(u,\tau)$$

where

$$\begin{aligned} \hat{S}_{G1}^{*}(u,\tau) &= \frac{1}{G}\sum_{g=1}^{G} \frac{1}{h_2} k_2\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right) w_g \frac{1}{N_g}\sum_{i=1}^{N_g}\phi_{ig}(u) \\ &= \frac{1}{G\bar{N}_G}\sum_{g=1}^{G}\sum_{i=1}^{N_g} \frac{1}{h_2} k_2\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right) w_g \frac{\bar{N}_G}{N_g}\phi_{ig}(u) \end{aligned}$$

and

$$R_{G1}(u,\tau) = \frac{1}{G}\sum_{g=1}^{G} \frac{1}{h_2} k_2\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right) w_g r_g(u)$$

where $\bar{N}_G = \frac{1}{G}\sum_{g=1}^{G} N_g$, $r_g(u)$ denotes the reminder term in Lemma 1, and

$$\begin{aligned} \phi_{ig}(u) &= e_1' J_g^{-1}(u)\left[1\left\{y_{ig} - z_{ig}'\alpha_g(u) < 0\right\} - u\right] z_{ig} \\ &= e_1' J_g^{-1}(u)\left[1\left\{\tilde{z}_{ig}'\gamma(U_{ig}) + x_g'\beta(U_{ig}, U_g) < z_{ig}'\alpha_g(u)\right\} - u\right] z_{ig}. \end{aligned}$$

First, note that

$$\|R_{G1}(u,\tau)\| \leq \max_g |r_g(u)| \frac{1}{G}\sum_{g=1}^{G}\left\| \frac{1}{h_2}k_2\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\right\|.$$

As in the proof of Lemma A3, we can show

$$\|R_{G1}\| \leq O_p(1) \times O_p\left(N_g^{-1}h_1^{-1}\ln N_g\right) = o_p\left(G^{-1/2}\right)$$

where we have made use of Assumption 9 and the result $\sup_g |r_g(u)| = O_p\left(N_g^{-1}h_1^{-1/2}\ln N_g\right)$ in Lemma 1.

Now we consider $\hat{S}_{G1}^{*}(u,\tau)$. Define

$$q_{ig}(u,\tau,\xi_{ig},U_{ig},U_g) = \frac{1}{\sqrt{G}}\frac{1}{\sqrt{\bar{N}_G}}\frac{\bar{N}_G}{N_g}k_2\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\phi_{ig}(u)$$

where $\xi_{ig} = (z_{ig}, x_g, w_g)$. Thus, we can write

$$\hat{S}_{G1}^{*}(u,\tau) = \frac{1}{\sqrt{G}\sqrt{\bar{N}_G}}\sum_{g=1}^{G}\sum_{i=1}^{N_g} q_{ig}(u,\tau,\xi_{ig},U_{ig},U_g).$$

Note that $q_{ig}(u,\tau,\xi_{ig},U_{ig},U_g)$ and $q_{jg}(u,\tau,\xi_{jg},U_{jg},U_g)$ for $i \neq j$ are not independent as both depend on U_g. We make use of the symmetrization technique commonly used in empirical process literature to show $\hat{S}_{G1}^{*}(u,\tau) = o_p\left(G^{-1/2}\right)$ uniformly in τ. Let $\left\{U_{ig}'\right\}$ be independent copies of $\left\{\xi_{ig}, U_{ig}, U_g\right\}$ on the same probability space, independent of $\{U_g\}$. Also on the same probability space let $\{\sigma_{ig}\}$ denote i.i.d. Rademacher random variables that are independent of $\left\{\xi_{ig}, U_{ig}, U_{ig}', U_g\right\}$. Next, Let E' and E_σ denote expectations taken with respect to $\left\{U_{ig}'\right\}$ and $\{\sigma_{ig}\}$ respectively. Following the arguments in the proof of Theorem 2.11.1 (van der

Vaart and Wellner, 1996), we can show that

$$
\begin{aligned}
& E\sup_{\tau}\left\|\sum_{g=1}^{G}\sum_{i=1}^{N_g} q(u,\tau,\xi_{ig},U_{ig},U_g)\right\| \\
= \; & E\sup_{\tau}\left\|\sum_{g=1}^{G}\sum_{i=1}^{N_g} q(u,\tau,\xi_{ig},U_{ig},U_g)-E'\left[q(u,\tau,\xi_{ig},U'_{ig},U_g)|\xi_{ig},U_{ig,}U_g,x_g,w_g\right]\right\| \\
\leq \; & EE'\sup_{\tau}\left\|\sum_{g=1}^{G}\sum_{i=1}^{N_g}\left(q(u,\tau,\xi_{ig},U_{ig},U_g)-q(u,\tau,\xi_{ig},U'_{ig},U_g)\right)\right\| \\
= \; & E_\sigma EE'\sup_{\tau}\left\|\frac{1}{2}\sum_{g=1}^{G}\sum_{i=1}^{N_g}\sigma_{ig}\left(2q(u,\tau,\xi_{ig},U_{ig},U_g)-2q(u,\tau,\xi_{ig},U'_{ig},U_g)\right)\right\| \\
\leq \; & \frac{1}{2}E_\sigma EE'\sup_{\tau}\left\|2\sum_{g=1}^{G}\sum_{i=1}^{N_g}\sigma_{ig}q(u,\tau,\xi_{ig},U_{ig},U_g)\right\|+\frac{1}{2}E_\sigma EE'\sup_{u,\tau}\left\|2\sum_{g=1}^{G}\sum_{i=1}^{N_g}\sigma_{ig}q(u,\tau,\xi_{ig},U'_{ig},U_g)\right\| \\
= \; & E_\sigma E\sup_{\tau}\left\|2\sum_{g=1}^{G}\sum_{i=1}^{N_g}\sigma_{ig}q(u,\tau,\xi_{ig},U_{ig},U_g)\right\| \\
\leq \; & C_M E\int_0^{\delta_n(\omega)}\sqrt{\log N(\varepsilon,A_n,\|\cdot\|}d\varepsilon \\
< \; & \infty
\end{aligned}
$$

where $\delta_n(\omega)=\sup_\tau\left|q(u,\tau,\xi_{ig}(\omega),U_{ig}(\omega),U_g(\omega))\right|$, A_n denotes the subset of R^n as the set of all vectors $\left\{q(u,\tau,\xi_{ig}(\omega),U_{ig}(\omega),U_g(\omega))\right\}$ for all $\omega\in\Omega$, and the last inequality follows since the class of functions $\mathcal{F}_q=\left\{q_{ig}(u,\tau,\xi_{ig},U_{ig},U_g)\colon\ \tau\in\mathcal{T}\right\}$ is Euclidean with a bounded envelop function. Consequently, we obtain $\hat{S}^{*}_{G1}(u,\tau)=O_p\left(G^{-1/2}\bar{N}_g^{-1}h_2^{-1}\right)$, and thus $\hat{S}_{G1}(u,\tau)=o_p\left(G^{-1/2}\right)$ uniformly over $\tau\in\mathcal{T}$.

Lemma A5: As N_G and G go to infinity, $\hat{S}_{G2}(u,\tau)=o_p(1)$.
Proof of Lemma A5: Recall

$$
\hat{S}_{G2}(u,\tau)=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}k'_2\left(\frac{\alpha_{g,1}(u)-x'_g\beta(u,\tau)}{h_2}\right)w_g\left(\hat{\alpha}_{g1}(u)-\alpha_{g,1}(u)\right)^2
$$

By Lemma 1, for each g, we have

$$\begin{aligned}
& \left(\hat{\alpha}_{g1}(u)-\alpha_{g,1}(u)\right)^2 \\
= \;& S_g^{*2}+2S_g^*\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)+\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)^2 \\
= \;& S_g^{*2}+O_p\left(N_g^{-1/2}\right)\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)+\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)^2
\end{aligned}$$

where $S_g^*=\frac{1}{N_g}\sum_{i=1}^{N_g}\phi_{ig}(u)$. Hence, by Assumptions 4, 8 and 9,

$$\begin{aligned}
\hat{S}_{G2}(u,\tau) &= S_{G2}^*(u,\tau)+O_p\left(h_2^{-2}\right)O_p\left(N_g^{-1/2}\right)\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)+\left(N_g^{-1}h_1^{-1/2}+h_1^{2s_1}\right)^2 \\
&= S_{G2}^*(u,\tau)+o_p\left(G^{-1/2}\right)
\end{aligned}$$

uniformly in τ, where

$$S_{G2}^*(u,\tau)=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_gS_g^{*2}.$$

Note that

$$\begin{aligned}
S_{G2}^* &= \frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\frac{1}{N_g^2}\sum_{i,j=1}^{N_g}\phi_{ig}\left(u\right)\phi_{jg}\left(u\right) \\
&= S_{G21}^*+S_{G22}^*,
\end{aligned}$$

where

$$S_{G21}^*=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{N_g^2}\sum_{i=1}^{N_g}\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\phi_{ig}^2\left(u\right)$$

and

$$S_{G22}^*=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\frac{1}{N_g^2}\sum_{i\neq j}^{N_g}\phi_{ig}\left(u\right)\phi_{jg}\left(u\right).$$

First, write S_{G21}^* as

$$S_{G21}^*=\frac{1}{\sqrt{G}\bar{N}_G^{3/2}h_2^2}\frac{1}{\sqrt{G\bar{N}_G}}\sum_{g=1}^{G}\sum_{i=1}^{N_g}\frac{\bar{N}_G^2}{N_g^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\phi_{ig}^2\left(u\right)$$

Then similar to (A12), we can show

$$\sqrt{G}\bar{N}_G^{3/2}h_2^2\left(S_{G21}^*-ES_{G21}^*\right)=O_p\left(1\right)$$

uniformly over $\tau \in \mathcal{T}$, and thus

$$S_{G21}^{*} = \left(\sqrt{G}\bar{N}_G^{3/2}h_2^2\right)^{-1}\left(ES_{G21}^{*} + O_p(1)\right) = o_p\left(G^{-1/2}\right)$$

uniformly over $\tau \in \mathcal{T}$.

Now we consider S_{G22}^{*}. Let

$$W_g(u) = \frac{1}{N_g^2}\sum_{i \neq j}^{N_g} \phi_{ig}(u)\,\phi_{jg}(u)$$

Note that

$$\begin{aligned} E\left[\phi_{ig}(u)\,\phi_{jg}(u)|U_{ig}, z_{ig}, \xi_{ig}\right] &= E\left[E\left[\phi_{ig}(u)\,\phi_{jg}(u)\,|U_{ig}, z_{ig}, U_g\right]|U_{ig}, z_{ig}\right] \\ &= E\phi_{ig}(u)\left[E\left[\phi_{jg}(u)\,|U_g\right]|U_{ig}, z_{ig}\right] = 0 \end{aligned}$$

thus, for a given g, $W_g(u)$ is a degenerate U-process for which the kernel function is of the VC type with an integrable envelope, then by Theorem 3.2 of Arcones and Giné (1994), we obtain,

$$\Pr\left(\left\|\frac{1}{N_g^2}\sum_{i \neq j}^{N_g}\phi_{ig}(u)\,\phi_{jg}(u)\right\| \geq tN_g^{-1}\ln^2 N_g\right) \leq c_1\exp(-c_2\ln^2 N_g) \leq N_G^{-3}$$

for large enough t. Therefore, we can deduce that there is a set Ω_G with $\Pr(\Omega_G) = O\left(N_G^{-2}\right)$, such that

$$\max_g\left\|\frac{1}{N_g^2}\sum_{i,j=1}^{N_g}\phi_{ig}(u)\,\phi_{jg}(u)\right\| = O\left(N_G^{-1}\ln^2 N_G\right).$$

In addition, similar to (A14), we can show that

$$\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}\left|k_2'\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right)\right| = O_p\left(h_2^{-1} + \left(Gh_2^3\right)^{-1/2}\ln^2 G\right).$$

Therefore,

$$\begin{aligned} \|S_{G22}^{*}\| &\leq \frac{1}{G}\sum_{g=1}^{G}\left\|\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right)w_g\right\| \times O\left(N_G^{-1}\ln^2 N_G\right) \\ &= O_p\left[h_2^{-1} + \left(Gh_2^3\right)^{-1/2}\ln^2 G\right] \times O\left(N_G^{-1}\ln^2 N_G\right) \\ &= o_p\left(G^{-1/2}\right). \end{aligned}$$

Collecting the above results establishes Lemma A5.

Proof of Theorem 2: We first establish the uniform consistency of $\hat{\beta}(u,\tau)$ over $\tau \in \mathcal{T}$ for a given $u \in \mathcal{U}$. From Lemma A1, we have

$$\hat{S}_G(u,\tau,b) = \frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\hat{\alpha}_{g,1}(u) - x_g' b}{h_2}\right) - \tau\right] w_g = S(u,\tau,b) + o_p(1)$$

uniformly over (τ, b). For a fixed τ, Assumptions 5 and 6 and standard argument in Newey and McFadden (1994) yields the consistency of $\hat{\beta}(u,\tau)$. From Lemma A.1 in Carroll et al. (1997), we can deduce that $\hat{\beta}(u,\tau) - \beta(u,\tau) = o_p(1)$ uniformly over $\mathcal{T}$. Next, a Taylor expansion of the estimating equation yields

$$\hat{S}_G(u,\tau) = \hat{Q}_{u,\tau}\left(\hat{\beta}(u,\tau) - \beta(u,\tau)\right)$$

uniformly over $\tau \in \mathcal{T}$ with probability approaching one, where

$$\hat{S}_G(u,\tau) = \frac{1}{G}\sum_{g=1}^{G}\left[K\left(\frac{\hat{\alpha}_{g1}(u) - x_g'\beta(u,\tau)}{h_2}\right) - \tau\right] w_g$$

and

$$\bar{Q}_{u,\tau} = \frac{1}{Gh_2}\sum_{g=1}^{G} k\left(\frac{\hat{\alpha}_{g1}(u) - x_g'\bar{\beta}(u,\tau)}{h_2}\right) w_g x_g'$$

with $\bar{\beta}(u,\tau)$ being on the line segment between $\hat{\beta}(u,\tau)$ and $\beta(u,\tau)$. The consistency $\hat{\beta}(u,\tau)$ and Lemma A2 yields $\bar{Q}_{u,\tau} = Q_{u,\tau} + o_p(1)$ uniformly over $\tau \in \mathcal{T}$, where the singular values of $Q_{u,\tau}$ are uniformly bounded away from zero in absolute values. Hence

$$\left(\hat{\beta}(u,\tau) - \beta(u,\tau)\right) = \left(Q_{u,\tau}Q_{u,\tau}'\right)^{-1} Q_{u,\tau}'\hat{S}_G(u,\tau) + o_p\left(\hat{S}_G(u,\tau)\right)$$

uniformly over $\tau \in \mathcal{T}$ with probability approaching one.

Now we consider $\hat{S}_G(u,\tau)$. By a Taylor expansion, we obtain

$$\hat{S}_G(u,\tau) = S_G(u,\tau) + \hat{S}_{G1}(u,\tau) + \hat{S}_{G2}(u,\tau) + \hat{S}_{G3}(u,\tau)$$

where

$$S_G(u,\tau) = \frac{1}{G}\sum_{g=1}^{G}\left[K_2\left(\frac{\alpha_{g1}(u) - x_g'\beta(u,\tau)}{h_2}\right) - \tau\right] w_g$$

$$\hat{S}_{G1}(u,\tau) = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2}k_2\left(\frac{\alpha_{g,1}(u) - x_g'\beta(u,\tau)}{h_2}\right) w_g\left(\hat{\alpha}_{g1}(u) - \alpha_{g,1}(u)\right)$$

$$\hat{S}_{G2}(u,\tau)=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^2}k_2'\left(\frac{\alpha_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\left(\hat{\alpha}_{g1}(u)-\alpha_{g,1}(u)\right)^2$$

and

$$\hat{S}_{G3}(u,\tau)=\frac{1}{G}\sum_{g=1}^{G}\frac{1}{h_2^3}k_2''\left(\frac{\bar{\alpha}_{g,1}(u)-x_g'\beta(u,\tau)}{h_2}\right)w_g\left(\hat{\alpha}_{g1}(u)-\alpha_{g,1}(u)\right)^3$$

where $\bar{\alpha}_{g,1}(u)$ is between $\hat{\alpha}_{g,1}(u)$ and $\alpha_{g,1}(u)$

By Lemma 1 and Assumption 9, we have

$$\hat{S}_{G3}(u,\tau)=O_p\left(\frac{\delta_G^3}{h_2^3}\right)=O_p\left(h_2^{-3}\right)O_p\left(N_G^{-3/2}+h_1^{3s_1}\right)=o_p\left(G^{-1/2}\right).$$

In addition, Lemmas A4 and A5 state that $\hat{S}_{G1}(u,\tau)=o_p\left(G^{-1/2}\right)$ and $\hat{S}_{G2}(u,\tau)=o_p\left(G^{-1/2}\right)$, and Lemma 3 states

$$S_G(u,\tau)=\frac{1}{G}\sum_{g=1}^{G}\left[1\left\{\alpha_{g1}(u)<x_g'\beta(u,\tau)\right\}-\tau\right]w_g+o_p\left(G^{-1/2}\right)$$

uniformly over $\tau\in\mathcal{T}$. Consequently, we obtain

$$\begin{aligned}&\sqrt{G}\left(\hat{\beta}(u,\tau)-\beta\left(u,\tau\right)\right)\\ =\;&\left(Q_{u,\tau}'Q_{u,\tau}\right)^{-1}Q_{u,\tau}'\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left(1\left\{\alpha_{g,1}(u)<x_g'\beta\left(u,\tau\right)\right\}-\tau\right)w_g+o_p\left(1\right)\end{aligned}$$

uniformly over $\tau\in\mathcal{T}$.

Define

$$\begin{aligned}f_{Gg}(\omega,\tau)&=\frac{1}{\sqrt{G}}\left(1\left\{\alpha_{g,1}(u)<x_g'\beta\left(u,\tau\right)\right\}-\tau\right)w_g\\&=\frac{1}{\sqrt{G}}\left(1\left\{y_{1g}(u)<z_{1g}'\alpha\left(u,\tau\right)\right\}-\tau\right)w_g\end{aligned}$$

$$\rho_G^2(\tau_1,\tau_2)=\sum_{g=1}^{G}E\left[f_{Gg}(\omega,\tau_1)-f_{Gg}(\omega,\tau_2)\right]^2$$

and

$$\rho^2(\tau_1,\tau_2)=(\tau_2-\tau_1)(1-\tau_2+\tau_1)Q_{ww}.$$

Then it is straightforward to verify the conditions of Theorem 10.6 in Pollard (1990) and we obtain

$$\frac{1}{\sqrt{G}}\sum_{g=1}^{G}\left(1\left\{y_{1g}(u) < x_g'\beta\left(u,\tau\right)\right\} - \tau\right)w_g \Longrightarrow \mathbb{G}\left(\cdot\right), \quad \text{in } \ l^{\infty}\left(\mathcal{T}\right)$$

where $\mathbb{G}\left(\cdot\right)$ is a zero-mean Gaussian process with uniformly continuous sample paths and covariance function $\Sigma(\tau,\tau') = \left[\min\left(\tau,\tau'\right) - \tau\tau'\right]Q_{ww}$. Finally, Theorem 2 follows from the Slustky Theorem.

Proof of Theorem 3: Lemma A2 and the consistency of $\hat{\beta}(u,\tau)$ imply that $\hat{Q}_{u,\tau}$ converges in probability uniformly over τ. In addition, by Assumption 7 we have $\hat{Q}_{ww} \rightarrow_p Q_{ww}$. Then, Theorem 3 follows easily.

7 References

Angrist, J.D. and J. Guryan (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements, The Economics of Education Review,27(5),483-503

Angrist, J. and Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. American Economic Review, 94, 1613–1634

Andrew Chesher (2003). Identification in Nonseparable Models,Econometrica , 71, 5,1405-1441

Arcones M.A. and E. Giné (1994). U-processes indexed by Vapnik-Červonenkis classes of functions with applications to asymptotics and bootstrap of U-statistics with estimated parameters, Stochastic Processes and their Applications, 52, 1, , 17-38

Arellano, M. and Bonhomme, S. 2016. Nonlinear panel data estimation via quantile regressions. Econometrics Journal 19, C61–C94.

Bousquet, O., 2002. A Bennett concentration inequality and its application to suprema of empirical processes. Comptes Rendus Mathematique, 334,495–500.

Carroll, R.J., J. Fan, I. Gijbels and M.P. Wand (1997). Generalized partially linear single-index models, Journal of the American Statistical Association, 92, 438, 477-489

Chen, S., 2018. "Sequential estimation of censored quantile regression models," Journal of Econometrics, Elsevier, vol. 207(1), pages 30-52.

Chernozhukov, V., and H. Christian (2006): Instrumental Quantile Regression Inference for Structural and Treatment Effect Models, Journal of Econometrics,132, 491-525.

Chernozhukov, V., and H. Christian (2008): Instrumental variable quantile regression: A robust inference approach, Journal of Econometrics, 142, 379-398

Chernozhukov, V., Chetverikov, D., and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. The Annals of Statistics, 42, 1564-1597.

Chetverikov, D. B. Larsen and C. Palmer (2016). IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade 84, 2, 809-833

Fernandes, M. G. Emmanuel, and H. Eduardo (2021). Smoothing Quantile Regressions, Journal of Business and Economic Statistics, 39, 1, 338-357Galvao, A. (2011). Quantile regression for dynamic panel data with fixed effects. Journal of Econometrics, 164, 142–157.

Galvao, A. and K. Kato (2016). Smoothed quantile regression for panel data, Journal of Econometrics, 193, 1, 92-112

Gu J. and V. Stanislav Volgushev (2019) Panel data quantile regression with grouped fixed effects, Journal of Econometrics, 213, 68-91

Hahn, J. and Meinecke, J. (2005). Time-invariant regressor in nonlinear panel model with fixed effects. Econometric Theory, 21, 455–469.

Hausman, J. and Taylor, W. (1981). Panel data and unobservable individual effects. Econometrica, 49, 1377–1398

He, X. K.M Tan and W. Zhou (2020), Smoothed quantile regression with large-scale inference, Journal of Econometrics, forthcoming.
Horowitz, J. (1998). Bootstrap Methods for Median Regression Models, Econometrica, 66, 6, 1327-1352
Jin, Z. and Ying, Z. (2004). Asymptotic theory in rank estimation for AFT model under fixed censorship. Parametric and Semiparametric Models with Applications to Reliability, Survival Analysis, and Quality of Life, Editors : M. Nikulin, N. Balakrishnan, M. Mesbah, and N. Limnios, Boston: Birkhauser, 107-120
Jun, S.J. (2009). Local Structural Quantile Effects in a Model with a Nonseparable Control Variable, Journal of Econometrics, 151, 82-97.
Kaplan, D. and Y. Sun (2017). Smoothed estimating equations for instrumental variables quantile regression, Econometric Theory, 33, 2017, 105–157
Kato, K., Galvao, A., and Montes-Rojas, G. (2012). Asymptotics for panel quantile regression models with individual effects. Journal of Econometrics, 170, 76–91.
Knight, K. (1998), Limiting distributions for L1 regression estimators under general conditions, Ann. Statist. 26(2): 755-770
Krueger, A.B. (1999). Experimental Estimates of Education Production Functions, The Quarterly Journal of Economics, Volume 114, Issue 2, May 1999, Pages 497–532,
Larsen, B. (2014). Occupational licensing and quality: Distributional and heterogeneous effects in the teaching profession. Working paper.
Newey, W.K. and D. McFadden, 1994, Large sample estimation and hypothesis testing, In: Engle, R.F. and McFadden, D., Eds., Handbook of Econometrics, Vol. 4, North Holland, Amsterdam.
Nolan D. and D. Pollard (1987). U-Processes: Rates of Convergence, Ann. Statist. 15 (2) 780 - 799,
Pakes A. and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," Econometrica, 57, 1027-1057.
Pollard D. (1990), Empirical Processes: Theory and Application, NSF-CBMS regional conference series in probability and statistics, Vol. 2, Institute of Mathematical Statistics and the American Statistical Association
Van der Vaart, A. and Wellner, J. (1996). Weak convergence and empirical processes. Springer Series in Statistics.

Table I: The Exogenous Case

Quantile	(G,N_G)	$(25,50)$		$(50,50)$		$(50,100)$		$(100,50)$	
(u,τ)	$\beta(u,\tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3,0.5)$	-1.049	-0.0453	0.6700	-0.0367	0.4740	-0.0107	0.4151	-0.0340	0.3249
$(0.3,0.5)$	-0.524	-0.0084	0.6358	-0.0008	0.4497	-0.0152	0.4044	0.0132	0.3038
$(0.3,0.7)$	0.000	0.0493	0.6797	0.0490	0.4858	0.0037	0.4228	0.0642	0.3542
$(0.5,0.3)$	0.524	-0.0322	0.6635	-0.0387	0.4819	-0.0192	0.4187	-0.0457	0.3130
$(0.5,0.5)$	0.000	-0.0097	0.6069	-0.0030	0.4361	-0.0112	0.3935	-0.0033	0.3009
$(0.5,0.7)$	0.524	0.0165	0.6479	0.0275	0.4862	-0.0070	0.4222	0.0411	0.3285
$(0.7,0.3)$	0.000	-0.0512	0.6860	-0.0333	1.0857	-0.0270	0.4182	-0.0580	0.3231
$(0.7,0.5)$	0.524	-0.0228	0.6261	-0.0069	0.4464	-0.0175	0.3997	-0.0207	0.3054
$(0.7,0.7))$	1.049	0.0120	0.6841	0.0362	0.4940	-0.0071	0.4287	0.0310	0.3319
	(G,N_G)	$(100,100)$		$(100,200)$		$(200,100)$		$(200,200)$	
(u,τ)	$\beta(u,\tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3,0.3)$	-1.049	-0.0202	0.2966	-0.0050	0.2807	-0.0296	0.2143	-0.0039	0.1930
$(0.3,0.5)$	-0.524	-0.0046	0.2881	-0.0071	0.2534	-0.0001	0.1998	0.0095	0.1847
$(0.3,0.7)$	0.000	0.0266	0.3380	0.0015	0.2684	0.0297	0.2394	0.0166	0.1942
$(0.5,0.3)$	-0.524	-0.0300	0.2931	-0.0084	0.2746	-0.0273	0.2033	-0.0086	0.1942
$(0.5,0.5)$	0.000	-0.0138	0.2874	-0.0041	0.2573	-0.0033	0.1978	0.0074	0.1847
$(0.5,0.7)$	0.524	0.0081	0.3102	-0.0030	0.2680	0.0153	0.2014	0.0122	0.1957
$(0.7,0.3)$	0.000	-0.0342	0.3093	-0.0095	0.2744	-0.0337	0.2060	-0.0104	0.1965
$(0.7,0.5)$	0.524	-0.0180	0.2847	-0.0099	0.2575	-0.0096	0.1997	0.0011	0.1895
$(0.7,0.7)$	1.049	0.0134	0.3062	-0.0083	0.2721	0.0129	0.2150	0.0129	0.1951

Table II: The Endogenous Case with $c_0 = 1$

Quantile	(G, N_G)	$(25, 50)$		$(50, 50)$		$(50, 100)$		$(100, 50)$	
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.5)$	-1.049	-0.172	0.855	-0.126	0.592	-0.068	0.491	-0.109	0.409
$(0.3, 0.5)$	-0.524	-0.082	0.819	-0.068	0.563	-0.060	0.482	-0.056	0.396
$(0.3, 0.7)$	0.000	0.001	0.846	-0.022	0.590	-0.053	0.527	-0.004	0.412
$(0.5, 0.3)$	-0.524	-0.161	0.820	-0.138	0.567	-0.067	0.476	-0.108	0.377
$(0.5, 0.5)$	0.000	-0.082	0.786	-0.087	0.546	-0.062	0.473	-0.077	0.377
$(0.5, 0.7)$	0.524	-0.022	0.835	-0.042	0.572	-0.059	0.512	-0.038	0.405
$(0.7, 0.3)$	0.000	-0.201	0.842	-0.150	0.580	-0.084	0.490	-0.138	0.406
$(0.7, 0.5)$	0.524	-0.097	0.819	-0.097	0.575	-0.067	0.483	-0.082	0.415
$(0.7, 0.7))$	1.049	-0.034	0.880	-0.055	0.605	-0.057	0.509	-0.043	0.443
	(G, N_G)	$(100, 100)$		$(100, 200)$		$(200, 100)$		$(200, 200)$	
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.3)$	-1.049	-0.0826	0.3250	-0.0693	0.3011	-0.0776	0.2381	-0.0587	0.2149
$(0.3, 0.5)$	-0.524	-0.0760	0.3237	-0.0657	0.2985	-0.0604	0.2272	-0.0530	0.2111
$(0.3, 0.7)$	0.000	-0.0605	0.3557	-0.0620	0.3203	-0.0420	0.2538	-0.0413	0.2279
$(0.5, 0.3)$	-0.524	-0.0971	0.3161	-0.0685	0.2932	-0.0818	0.2240	-0.0579	0.2059
$(0.5, 0.5)$	0.000	-0.0806	0.3180	-0.0674	0.2925	-0.0700	0.2199	-0.0519	0.2035
$(0.5, 0.7)$	0.524	-0.0611	0.3469	-0.0598	0.3208	-0.0580	0.2503	-0.0481	0.2223
$(0.7, 0.3)$	0.000	-0.1031	0.3202	-0.0728	0.3001	-0.0885	0.2316	-0.0635	0.2115
$(0.7, 0.5)$	0.524	-0.0867	0.3225	-0.0681	0.2962	-0.0748	0.2287	-0.0536	0.2080
$(0.7, 0.7)$	1.049	-0.0779	0.3537	-0.0627	0.3336	-0.0599	0.2523	-0.0509	0.2252

Table III: The Endogenous Case with $c_0 = 0.5$

Quantile	(G, N_G)	$(25, 50)$		$(50, 50)$		$(50, 100)$		$(100, 50)$	
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.5)$	-1.049	-0.1097	0.8683	-0.0830	0.6039	-0.0359	0.5064	-0.0885	0.4241
$(0.3, 0.5)$	-0.524	-0.1155	0.8471	-0.0871	0.5810	-0.0788	0.5001	-0.0673	0.4112
$(0.3, 0.7)$	0.000	-0.0897	0.8928	-0.0779	0.6222	-0.1049	0.5558	-0.0397	0.4288
$(0.5, 0.3)$	-0.524	-0.1157	0.8427	-0.1065	0.5800	-0.0403	0.4925	-0.0837	0.3920
$(0.5, 0.5)$	0.000	-0.1119	0.8107	-0.1080	0.5718	-0.0828	0.4966	-0.0906	0.3906
$(0.5, 0.7)$	0.524	-0.1086	0.8820	-0.0989	0.5924	-0.1108	0.5378	-0.0702	0.4197
$(0.7, 0.3)$	0.000	-0.1414	0.8621	-0.1161	0.6042	-0.0557	0.5060	-0.1157	0.4195
$(0.7, 0.5)$	0.524	-0.1285	0.8521	-0.1181	0.6000	-0.0903	0.5098	-0.0964	0.4306
$(0.7, 0.7))$	1.049	-0.1193	0.9194	-0.1157	0.6280	-0.1088	0.5366	-0.0773	0.4618
	(G, N_G)		100,100		100,200		200,100		200,200
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.3)$	-1.049	-0.0641	0.3346	-0.0573	0.3112	-0.0670	0.2453	-0.0517	0.2216
$(0.3, 0.5)$	-0.524	-0.0892	0.3357	-0.0818	0.3099	-0.0681	0.2337	-0.0614	0.2169
$(0.3, 0.7)$	0.000	-0.0914	0.3719	-0.0926	0.3345	-0.0617	0.2618	-0.0593	0.2355
$(0.5, 0.3)$	-0.524	-0.0835	0.3288	-0.0577	0.3049	-0.0732	0.2317	-0.0500	0.2125
$(0.5, 0.5)$	0.000	-0.0956	0.3316	-0.0828	0.3057	-0.0793	0.2268	-0.0614	0.2095
$(0.5, 0.7)$	0.524	-0.0925	0.3611	-0.0905	0.3370	-0.0781	0.2600	-0.0670	0.2303
$(0.7, 0.3)$	0.000	-0.0878	0.3335	-0.0591	0.3046	-0.0790	0.2388	-0.0562	0.2177
$(0.7, 0.5)$	0.524	-0.0997	0.3345	-0.0842	0.3101	-0.0828	0.2350	-0.0618	0.2149
$(0.7, 0.7)$	1.049	-0.1108	0.3711	-0.0928	0.3424	-0.0776	0.2624	-0.0711	0.2319

Table IV: The Endogenous Case with $c_0 = 1.5$

Quantile	(G, N_G)	$(25, 50)$		$(50, 50)$		$(50, 100)$		$(100, 50)$	
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.5)$	-1.049	-0.2765	0.8581	-0.1959	0.5900	-0.1183	0.4911	-0.1461	0.4017
$(0.3, 0.5)$	-0.524	-0.0478	0.8012	-0.0436	0.5467	-0.0348	0.4684	-0.0419	0.3856
$(0.3, 0.7)$	0.000	0.1169	0.8221	0.0648	0.5667	0.0210	0.5039	0.0477	0.4003
$(0.5, 0.3)$	0.524	-0.2491	0.8286	-0.1963	0.5631	-0.1139	0.4748	-0.1455	0.3711
$(0.5, 0.5)$	0.000	-0.0489	0.7694	-0.0628	0.5305	-0.0355	0.4568	-0.0596	0.3667
$(0.5, 0.7)$	0.524	0.0994	0.8034	0.0404	0.5549	0.0129	0.4905	0.0141	0.3914
$(0.7, 0.3)$	0.000	-0.2906	0.8528	-0.2086	0.5727	-0.1303	0.4855	-0.1754	0.4011
$(0.7, 0.5)$	0.524	-0.0605	0.7970	-0.0722	0.5590	-0.0401	0.4642	-0.0640	0.4051
$(0.7, 0.7))$	1.049	0.0904	0.8423	0.0318	0.5866	0.0146	0.4889	0.0128	0.4299
	(G, N_G)	$(100, 100)$		$(100, 200)$		$(200, 100)$		$(200, 200)$	
(u, τ)	$\beta(u, \tau)$	Bias	SD	Bias	SD	Bias	SD	Bias	SD
$(0.3, 0.3)$	-1.049	-0.1130	0.3233	-0.0917	0.2975	-0.0951	0.2344	-0.0730	0.2111
$(0.3, 0.5)$	-0.524	-0.0577	0.3146	-0.0475	0.2900	-0.0499	0.2229	-0.0407	0.2068
$(0.3, 0.7)$	0.000	-0.0119	0.3436	-0.0175	0.3092	-0.0110	0.2476	-0.0138	0.2225
$(0.5, 0.3)$	-0.524	-0.1241	0.3119	-0.0902	0.2887	-0.0980	0.2200	-0.0718	0.2028
$(0.5, 0.5)$	0.000	-0.0614	0.3090	-0.0486	0.2833	-0.0582	0.2161	-0.0396	0.1991
$(0.5, 0.7)$	0.524	-0.0148	0.3359	-0.0158	0.3094	-0.0274	0.2421	-0.0193	0.2164
$(0.7, 0.3)$	0.000	-0.1307	0.3147	-0.0968	0.3016	-0.1052	0.2280	-0.0775	0.2086
$(0.7, 0.5)$	0.524	-0.0691	0.3134	-0.0482	0.2865	-0.0631	0.2249	-0.0416	0.2030
$(0.7, 0.7)$	1.049	-0.0290	0.3400	-0.0179	0.3254	-0.0298	0.2447	-0.0214	0.2194