Far East Journal of Mathematical Sciences (FJMS)
© 2018 Pushpa Publishing House, Allahabad, India
http://www.pphmj.com
http://dx.doi.org/10.17654/MS103020451
Volume 103, Number 2, 2018, Pages 451-465 **ISSN: 0972-0871**

ESTIMATING QUANTILE REGRESSION FOR DYNAMIC PANEL DATA USING TWO-STAGE APPROACH

A. R. Elhoussainy[1]**, A. E. Ahmed**[1] **and M. A. Hossameldin**[2]

[1]Department of Applied Statistics and Econometrics
Institute of Statistical Studies and Research (ISSR)
Cairo University
Egypt
e-mail: aham103@yahoo.com

[2]Department of Economics
Faculty of Business Administration, Economics
and Political Science (BAEPS)
The British University in Egypt
Egypt
e-mail: hossameldin.ahmed90@gmail.com

Abstract

There are two main approaches in handling the problem of endogeneity in quantile regression for dynamic panel data. The widely used approach considers the inclusion of reliable instrumental variable(s) to minimize the biasedness problem. Our new approach is concerned with the implication of two-stage quantile regression for dynamic panel data. In order to gain the privilege of robustness for the proposed estimation process, both stages will be estimated by quantile

Received: September 9, 2017; Accepted: October 27, 2017

2010 Mathematics Subject Classification: 62P20, 62G08.

Keywords and phrases: endogeneity, dynamic panel data, two-stage quantile regression, quantile regression.

regression with the same level of quantile. We provide the derivation of asymptotic distribution for the new proposed structural parameter. In addition, we obtain the general form of variance-covariance matrix of the proposed estimator and prove its consistency.

1. Introduction

Estimating dynamic panel data (DPD) models through quantile regression (QR) approach has been increasingly considered in the literature over the last decade. It is a type of models which capture the dynamic effect through the inclusion of lagged value(s) for the response variable. As a result of that, it allows defining the dynamic effect for both dimensions time and cross sections (Moon and Weidner [20]).

Unlike some conventional estimation methods, QR does not depend on explaining the mean relationship but it presents an overview on any pre-determined level of quantile (Gu and Koenker [11]). Several studies have considered the approach of estimation of quantile regression in dynamic panel data (QRDPD) models. The presentation of QRDPD has essentially discussed from the perspective of handling the arising problem of endogeneity. Due to the involvement of regressors of lagged dependent variables, the correlation between them and innovation term generate biased estimators [3].

The first approach relies on adding instrumental variable(s) that is/are not correlated with innovations which is similar to the estimation framework proposed by Chernozhukov and Hansen [8]. Galvao [10] utilized the process of Chernozhukov and Hansen [8] along with lagged instruments to minimize the bias in DPD with fixed effects (FEs) models.

The old version of TSQR approach has been addressed by Amemiya [1], which outlined double-stage least absolute deviation (DSLAD) estimation of structural equations' parameters. Powell [23] provided the asymptotic distribution of DSLAD structural parameter. Both approaches by Amemiya and Powell described the median relationship, however, least absolute deviation (LAD) is not the main interest for many researchers. Chen and Portnoy [7] estimated TSQR when the first stage is based on trimmed least square (TLS) and LAD with symmetric error term for both stages. Other empirical studies are provided by Kemp [12] who used lease absolute error difference (LAED) in both stages. Another approach by MaCurdy and Timmins [19] presented estimation procedure for QR model with ARMA structure. Kim and Muller [16] proposed TSQR where both stages are estimated by QR, along with the asymptotic representation for the structural parameter. Variance reduction for QR models with endogeneity, asymptotic distribution of TSQR and inconsistency transmission property is also provided by Kim and Muller [14], where the first stage is estimated by DSLAD and TSLS.

We are concerned with the implication of *the second approach* of handling endogeneity in two-stage quantile regression for dynamic panel data (TSQRDPD) model. TSQRDPD works mainly on estimating the endogenous variable (lagged dependent variable) by its lagged values in the first stage. Afterwards, it plugs the resulted new estimated values in the original model as the second stage of estimation.

2. Estimation of TSQRDPD Model

TSQRDPD provides a much wider scope for interpreting the results at several quantiles of variables of study. Hence, the parameters along with their estimators are evaluated at different quantiles, therefore, not the mean relationships are the presented but larger scale is provided according to each quantile level. The main aim is to estimate the structural parameter $\omega_0 = (\Gamma_0', \Lambda_0')$ which is presented as follows:

$$y_{it} = Y_{i,t-1}\Gamma_0 + X_{1it}\Lambda_0 + U_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{1}$$

It can also be represented in compressed form as follows:

$$y_{it} = Z_{it}\omega_0 + U_{it},$$
$$U_{it} = \mu_i + \varepsilon_{it}, \quad i = 1, \ldots, N, \quad t = 2, \ldots, T. \tag{2}$$

For *i*th cross section and *t*th time period, $[y_{it}, Y_{i,t-1}]$ is an $NT \times (\mathcal{W} + 1)$ matrix of endogenous variables, y_{it} is the variable of study, $Y_{i,t-1}$ is the first lagged value of the dependent variable, X'_{1it} is an $NT \times \mathcal{L}_1$ of exogenous variables, the set of explanatory variables that are included in the model $Z_{it} = [y_{i,t-1}, X_{1,it}]$, the structural parameter $\omega_0 = (\Gamma'_0, \Lambda'_0)'$, ε_{it} is an $NT \times 1$ vector, μ_i is a vector of $I_N \times 1$. X_{2it} is an $NT \times \mathcal{L}_2$ exogenous variables that are not involved in the model in (1) and used to estimate $Y_{i,t-1}$. It should be taken into consideration that $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$. X_{2it} is an instrumental variable that can be used in the first stage to estimate the endogenous variable $Y_{i,t-1}$. According to Ma and Koenker [18], $Q_\theta(.)$ is the quantile function at (θ) conditional on Y_{it}. The problem of endogeneity would exist if $Q_\theta(U_{it} | Y_{it}) \neq Q_\theta(U_{it})$. This inequality between $Q_\theta(U_{it} | Y_{it})$ and $Q_\theta(U_{it})$ can represent an explanation of endogeneity in quantile regression. τ and θ are quantile levels for the first and the second stages, respectively.

The first stage is to estimate the endogenous variable Y_{it-1}. By collecting the observations over time periods and cross sections, the first stage model can be proposed in the matrix form as follows:

$$Y_{it-1} = X_{it}\Phi_0 + V_{it-1}, \quad i = 1, ..., N, \quad t = 2, ..., T. \tag{3}$$

The model in (3) is considered to be specified correctly, where $Y_{i,t-1}$ is the lagged value of dependent variable and $X_{it} = [X_{1it}, X_{2it}]$ which is matrix of size $NT \times \mathcal{L}$. Φ_0 is matrix of the first stage parameters that are unknown of size $\mathcal{L} \times \mathcal{W}$. V_{it} is an $NT \times \mathcal{W}$ matrix of observations of error term over N and T.

The second stage is to estimate the dependent variable y_{it}. By aggregating the observations over time periods and cross sections, the second stage model can be proposed as follows:

$$y_{it} = X_{it}\varphi_0 + v_{it}, \tag{4}$$

where $\varphi_0 = \begin{bmatrix} \Phi_0, & \begin{matrix} I_{\mathcal{L}_1} \\ 0 \end{matrix} \end{bmatrix} \omega_0 = \mathbb{N}(\Phi_0)\omega_0$ and the error term in (4) can be modeled as follows: $v_{it} = V_{it}\Gamma_0 + U_{it}$. Φ_0 which is obtained from the first stage will be included in φ_0 which represents the structural parameter in the second stage as shown in (4). The TSQRDPD estimator $\hat{\omega}_0 = (\hat{\Gamma}_0', \hat{\Lambda}_0')'$ for ω_0 can be obtained by solving the following minimization argument:

$$\min_{\omega_0} \mathfrak{D}_{NT}(\omega_0, \hat{\varphi}_0, \hat{\Phi}_0, \theta, \tau)$$

$$= \min_{\omega_0} \sum_{i=1}^{N}\sum_{t=1}^{T} \rho_\theta(\tau y_{it} + (1-\tau)x_{it}'\hat{\varphi}_0 - x_{it}'\mathbb{N}(\hat{\Phi}_0)\omega_0), \tag{5}$$

where y_{it} and x_{it}' are the observations of the dependent and independent variables, respectively, for ith cross section for tth time period. Both θ and τ are quantile levels for the first and second stages, respectively. They range from 0 to 1. The orders of quantiles are assumed to take the same value for consistency purposes. $\rho_\theta(g) = g\psi_\theta(g)$, where $\psi_\theta(g) = \theta - I_{(g \leq 0)}$ and $I(.)$ is the indicator function defined by Koenker and Bassett [17].

The formulation of the dependent variable in two stages is $\tau y_{it} + (1-\tau)x_{it}'\hat{\varphi}_0$. The formulation in (5) can be partitioned into two minimization arguments to obtain the two stages' estimators $\hat{\Phi}$ and $\hat{\varphi}$ based on the generalization provided by Amemiya [1] as follows:

$$\min_{\varphi} \sum_{i=1}^{N}\sum_{t=1}^{T} \rho_\theta(y_{it} - x_{it}\varphi)$$

and

$$\min_{\Phi_r} \sum_{i=1}^{N}\sum_{t=1}^{T} \rho_\theta(y_{rit} - x_{it}\Phi_r), \quad r = 1, 2, ..., \mathcal{W}, \tag{6}$$

where φ and Φ_r are $\mathcal{W} \times 1$ vectors, Y_{rth} is rth in (i, t)th elements of Y, Φ_r is the rth element in Φ which is obtained in the first stage of the estimation for the correspondent estimated endogenous variable. The adopted value(s) of quantile orders (τ) which determines the estimator in the first stage should be chosen carefully, the invariance property of QR can be utilized to show that TSQRDPD estimator does not vary according to the quantile order in the first stage.

3. Asymptotic Normality of TSQRDPD (θ, τ)

This section is devoted to show the asymptotic normality of TSQRDPD (θ, τ) attained by the derivation of its asymptotic illustration. It should be taken into consideration that the robustness property defined by Koenker and Bassett [17] can be evidently captured from the asymptotic illustration of TSQRDPD (θ, τ) because it relies on bounded functions of ρ_θ and ρ_τ. TSQRDPD (θ, τ) would not conserve its privilege of being robust to outliers if the first stage does not rely on an estimation method that is robust across the varying levels of quantiles. The distribution of TSQRDPD parameter can be presented as follows:

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_0) \xrightarrow{d} N(0, \Re\zeta\Re'),$$

where

$$\Re = Q_{zz}^{-1}\mathbb{N}(\Phi_0)'[I, -Q^* Q_1^{-1}\Gamma_{01}, -Q^* Q_2^{-1}\Gamma_{02}, ..., -Q^* Q_{\mathcal{W}}^{-1}\Gamma_{0\mathcal{W}}],$$

$$Q_{zz} = \mathbb{N}(\Phi_0)' Q^* \mathbb{N}(\Phi_0), \quad Q^* = E\{f(0 | x_{it}) x_{it} x'_{it}\}$$

and

$$Q_r = E\{g_r(0 | x_{it}) x_{it} x'_{it}\},$$

where $r = 1, 2, ..., \mathcal{W}$, $\hat{\zeta} = (NT)^{\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\Delta} x_{it} x'_{it}$, $\hat{\Delta} = \psi_\theta(\hat{\mathbb{K}}_{rit})\psi_\theta(\hat{\mathbb{K}}_{r^*it})$, where $(\hat{\mathbb{K}}_{rit}) = v_{it}$ and $(\hat{\mathbb{K}}_{r^*it}) = V_{r-1, it-1}$, where r would be $2 \le r \le (\mathcal{W} + 1)$.

3.1. Assumptions

Acquiring the asymptotic illustration of asymptotic normality for TSQRDPD (θ, τ) would require setting several conditions according to Pollard [21] and Powell [23]. The assumptions are as follows:

Assumption 1. The sequence $\{(x_{it}, v_{it}, V_{it-1})\}$ is independent and identically distributed (i.i.d) over *i*th cross section for *t*th time period.

Assumption 2. $E\| x_{it} \|^3 < \infty,$ where $\| . \|$ is a Euclidean norm. This assumption is the generalization of the outcomes in Powell [23] for the process $S_{NT}(.)$ to follow the stochastic equicontinuity. Its significance stems from handling the weak convergence of non-differentiable objective functions to obtain the asymptotes of estimators when both N and T tend to ∞.

Assumption 3. $\mathbb{N}(\Phi_0)$ is a matrix of full rank. This assumption is abstracted conventionally in the literature to guarantee that the estimation would be possible, especially when linear programming softwares are used.

Assumption 4. The conditional destiny functions for v_{it} and V_{rit-1} are $f(.|x_{it})$ and $g_r(.|x_{it}),$ respectively, which are continuous for all the values x across the panels, where $r = 1, 2, ..., \mathcal{W}.$ In addition, the matrices $Q^* = E\{f(0|x_{it})x_{it}x'_{it}\}$ and $Q_r = E\{g_r(0|x_{it})x_{it}x'_{it}\}$ are positive definite and this assumption is set in the literature as one of the basic ones.

Assumption 5. $E\{\psi_\theta(v_{it})|x_{it}\} = E\{\psi_\theta(V_{rit-1})|x_{it}\} = 0.$ The essentiality of this assumption is to guarantee that in the presence of an intercept term, it becomes normal at a particular level of θ of the distribution of both v_{it} and V_{it-1}.

Lemma 1. *The presentation of the stochastic process* $S_{NT}(.)$ (*defined in Assumption* 2) *suggested by Andrews* [2] *in the frame of TSQRDPD can be proposed as follows*:

$$\underset{\| \mathfrak{B} \| \leq C}{Sup} \| S_{NT}(\mathfrak{B}) - S_{NT}(0) + \tau^{-1}Q^*\mathfrak{B} \| = o_p(1).$$

Proof. We redefine Andrews [2] empirical process $S_{NT}(.)$ for TSQRDPD as follows:

$$S_{NT}(\mathfrak{B}) \equiv (NT)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} s(\Psi_{it}, \mathfrak{B}), \tag{7}$$

where $\dim(\mathfrak{B}) = \mathcal{L}_1 + \mathcal{L}_2$ is a vector of unknown structural parameter of dimension $\mathcal{L} \times 1$, $\Psi_{it} = (v_{it}, x'_{it})'$, ψ_θ is the function defined by Koenker and Bassett [17]. Also, $s(\Psi_{it}\mathfrak{B}) = x_{it}\psi_\theta(\tau v_{it} - NT^{-1/2}x'_{it}\mathfrak{B})$, hence (7) can be put in the following form:

$$S_{NT}(\mathfrak{B}) \equiv (NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} s(\Psi_{it}, \mathfrak{B})$$

$$\equiv (NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} x_{it}\psi_\theta(\tau v_{it} - NT^{-1/2}x'_{it}\mathfrak{B}). \tag{8}$$

Based on Powell [23], we reformulate the structural parameter of interest $\mathfrak{B}$ as follows:

$$\hat{\mathfrak{B}} \equiv \sqrt{NT}[\mathbb{N}(\hat{\Phi}_0)(\hat{\omega}_0 - \omega_0) - (1-\tau)(\hat{\varphi}_0 - \varphi_0) + (\hat{\Phi}_0 - \Phi_0)\Gamma_0]. \tag{9}$$

According to Assumption 3 and Kim and Muller [16], we obtain

$$\underset{\|\mathfrak{B}\| \le C}{Sup} \| S_{NT}(\mathfrak{B}) - S_{NT}(0) + \tau^{-1}Q^*\mathfrak{B} \| = o_p(1),$$

$$S_{NT}(0) = (NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} x_{it}\psi_\theta(\tau v_{it}) = O_p(1), \tag{10}$$

where C is a finite positive scalar, and $o_p(1)$ denotes that the stochastic sequence would converge in probability to zero when both dimensions of N and T tend to ∞. The result in (10) is the generalization of Lemma 1 proved by Powell [23] when x_{it} is random that follows the special case of Lemma 4.1 by Bickel [4].

Lemma 2. *The convergence of the process defined in Assumption* 2 *can be proposed as follows*:

$$\underset{\|\mathfrak{B}_1-\mathfrak{B}_2\|\le C}{Sup} \| M_{NT}(\mathfrak{B}_1) - M_{NT}(\mathfrak{B}_2) \| = o_p(1).$$

Proof. By applying Andrews [2] process in our frame of TSQRDPD, process $M_{NT}(\mathfrak{B})$ can be defined as follows:

$$M_{NT}(\mathfrak{B}) \equiv \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} [s(\Psi_{it}\mathfrak{B}) - E\{s(\Psi_{it}\mathfrak{B})\}]$$

$$= S_{NT}(\mathfrak{B}) - E(S_{NT}(\mathfrak{B})). \tag{11}$$

The result obtained in (10) can be similarly applied on (11) but the following two conditions set by Andrews should be satisfied:

Condition 1. $s(., \mathfrak{B})$ follows entropy condition by Pollard [21] with envelop function of $\tilde{S}(\Psi_{it})$. The quantification of the distance can be determined by using a measurable envelop function F satisfying for each f, $| f | \le F$.

Condition 2. $\lim_{N,T\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E[\{\tilde{S}(\Psi_{it})\}^{\delta}] < \infty$, where δ takes the same value of the elements of the set $\{\tilde{S}(\Psi_{it})\}$, where $\delta \ge 2$. Based on the above two conditions along with Theorem (2.1) in Andrews [2], the term in (11) will result in

$$\underset{\|\mathfrak{B}_1-\mathfrak{B}_2\|\le C}{Sup} \| M_{NT}(\mathfrak{B}_1) - M_{NT}(\mathfrak{B}_2) \| = o_p(1). \tag{12}$$

By letting type I function to be $f_1(., \mathfrak{B}) = f_1(\Psi_{it}, \mathfrak{B}) = x$ and type II function to be $f_2(., \mathfrak{B}) = \psi_\theta(\tau v_{it} - NT^{-1/2} x'_{it}\mathfrak{B})$, $C > 0$ and their Lipschitz coefficient is $(\mathfrak{B}) = \| x_{it} \|$. As a results of that, $s(., \mathfrak{B})$ in condition 1 can be viewed as a the result of product of $f_1(., \mathfrak{B})$ and $f_2(., \mathfrak{B})$, hence, it follows Pollard's entropy condition as shown in condition 1 and its envelope is $\max(1, \| x_{it} \|)$. Condition 2 can be also be verified to be represented as

$$\lim_{N,T\to\infty} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E[\{\widetilde{S}(\Psi_{it})\}^{\delta}] = E[\{\max(1, \| x_{it} \|)\}^{\delta}]$$

which is considered to the generalization of Assumption 2.

Lemma 3. *Based on Lemmas* 1 *and* 2, *the distribution of the structural parameter can be presented as follows, and subsequently its distribution can be proposed*

$$\underset{\| E \| \le c^*}{Sup} \| S_{NT}^*(E) - E(S_{NT}^*(0)) - \tau^{-1} Q^* E \| = o_p(1).$$

Proof. $\mathfrak{B}_1, \mathfrak{B}_2$ are defined in R^K and $c^* > 0$ is defined in R^+, and by defining the following terms: $\mathfrak{B}_1 = (NT)^{-1/2} E_1, \mathfrak{B}_2 = (NT)^{-1/2} E_2$, hence the following representations can be illustrated in a similar way as in (12) as follows:

$$M_{NT}^*(E) = S_{NT}^*(E) - E(S_{NT}^*(E)). \tag{13}$$

$S_{NT}^*(E)$ is similarly defined as in (7) and (8) as follows:

$$S_{NT}^*(E) = (NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} s(\Psi_{it}, E)$$

$$= (NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} x_{it} \psi_{\theta}(\tau v_{it} - (NT)^{-1/2} x_{it}' E). \tag{14}$$

Therefore, it would be clear to similarly prove as in (12) that

$$\underset{\| E_1 - E_2 \| \le c^*}{Sup} \| M_{NT}(E_1) - M_{NT}(E_2) \| = o_p(1). \tag{15}$$

From (13) and by substituting $E_1 = E$ and $E_2 = 0$ in (15), we can obtain following:

$$\underset{\| E \| \le c^*}{Sup} \| S_{NT}^*(E) - E(S_{NT}^*(0)) - \{E(S_{NT}^*(E)) - E(S_{NT}^*(0)) \| = o_p(1), \tag{16}$$

where $S_{NT}^*(0) = (NT)^{-\frac{1}{2}} x_{it} \psi_{\theta}(v_{it}) = O_p(1)$ as it converges in probability to zero in distribution under Assumptions 1-4 and Lindeberg-Levy central limit theorem (CLT).

It should be noted that the term $E(S_{NT}^{*}(E)) - E(S_{NT}^{*}(0))$ tends to be equal to $(-\tau^{-1}Q^{*}E)$ as proved - in details - by Kim and Muller [15]. Q^{*} is defined by Assumption 4. By utilizing the central limit theorem, therefore, the term in (16) can be reformulated as follows:

$$\underset{\| E \| \leq c^{*}}{Sup} \| S_{NT}^{*}(E) - E(S_{NT}^{*}(0)) - \tau^{-1}Q^{*}E \| = o_{p}(1). \tag{17}$$

The argument in (16) will be the initiation of next stage of the proof to provide the estimator of first stage for TSQRDPD. The process of $S_{NT}^{*}(\hat{E})$ can be reformulated as follows:

$$S_{NT}^{*}(\hat{E}) = S_{NT}^{*}(0) - \tau^{-1}Q^{*}\hat{E} + O_{p}(1) = O_{p}(1). \tag{18}$$

Corollary 1. *According to Assumptions* 1 *and* 2, *it is easy to prove the following*:

- $\hat{E} = (\tau - 1)\sqrt{NT}(\hat{\varphi} - \varphi_{0}) + \sqrt{NT}(\hat{\Phi} - \Phi_{0})\Gamma_{0}$.
- *The distribution of the first stage parameter follows:* $\sqrt{NT}(\hat{\Phi} - \Phi_{0}) = O_{p}(1)$ *and the distribution for the second stage structural parameter follows*: $\sqrt{NT}(\hat{\varphi} - \varphi_{0}) = O_{p}(1)$

and subsequently, the distribution of structural parameter would converge as $\sqrt{NT}(\hat{\Phi} - \Phi_{0}) = O_{p}(1)$ *that is based on Lemma* (A.2) *of Carroll and Ruppert* [6].

Lemma 4. *The distribution of the structural parameter in* (1) *and* (2) *can be defined as follows*:

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_{0}) \xrightarrow{d} N(0, \mathfrak{R}\zeta\mathfrak{R}').$$

Proof. By plugging the results obtained in Lemmas 2 and 3 and considering Lemmas (A.8)-(A.12) by Powell [23] in double least absolute deviation (DLAD), we now formulate a generalization of these results to establish our primary asymptotic illustration of TSQRDPD (θ, τ) as follows:

$$(NT)^{1/2}(\hat{\omega} - \omega_0) = \tau Q_{zz}^{-1}\mathbb{N}(\Phi_0)' S_{NT}^{*}(\hat{E}) + o_p(1), \tag{19}$$

where Q_{zz} represents the term $\mathbb{N}(\Phi_0)Q^{*}\mathbb{N}(\Phi_0)'$. We reformulate the asymptotic representation in (19) to provide our asymptotic illustration as follows:

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_0)$$
$$= \tau Q_{zz}^{-1}\mathbb{N}(\Phi_0)'\Bigg[(NT)^{\frac{1}{2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\tau x_{it}\psi_{\theta}(v_{it})$$
$$+ (1-\tau)Q^{*}(NT)^{\frac{1}{2}}(\hat{\varphi} - \varphi_0) - Q^{*}(NT)^{\frac{1}{2}}(\hat{\Phi} - \Phi_0)\Gamma_0\Bigg] + o_p(1). \tag{20}$$

Depending on the results of (A.12) in Powell [23], and (A.9) in Kim and Muller [13] regarding $(NT)^{\frac{1}{2}}(\hat{\varphi} - \varphi_0)$ and $(NT)^{\frac{1}{2}}(\hat{\Phi} - \Phi_0)$, we can reevaluate them in our new frame of TSQRDPD as follows:

$$(NT)^{\frac{1}{2}}(\hat{\varphi} - \varphi_0) = Q^{*}(NT)^{-\frac{1}{2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\tau x_{it}\psi_{\theta}(v_{it}) + o_p(1),$$
$$(NT)^{\frac{1}{2}}(\hat{\Phi} - \Phi_0) = Q_r^{-1}(NT)^{-\frac{1}{2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\tau x_{it}\psi_{\theta}(v_{it-1}) + o_p(1), \tag{21}$$

where Q^{*} and Q_r are defined in Assumption 4. By plugging (20) in (21) and after running few calculations and simplifications, we can propose our final asymptotic illustration for TSQRDPD (θ, τ):

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_0)$$
$$= (NT)^{-\frac{1}{2}}\sum_{i=1}^{N}\sum_{t=1}^{T}Q_{zz}^{-1}\mathbb{N}(\Phi_0)' x_{it}\psi_{\theta}(v_{it})$$
$$- (NT)^{-\frac{1}{2}}\sum_{r=1}^{\mathcal{W}}\sum_{i=1}^{N}\sum_{t=1}^{T}Q_{zz}^{-1}\mathbb{N}(\Phi_0)' Q^{*}Q_r^{-1}\Gamma_{0r}(V_{rit-1}) + o_p(1). \tag{22}$$

Summarizing, the result in (22) can be rewritten as follows:

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_0) = \Re(NT)^{-\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{W}_{it} + o_p(1). \tag{23}$$

Recall Assumption 5. It is clear to verify that $E(\mathbb{W}_{it}) = 0$ as long as the same level of quantile is used in the two stages of TSQRDPD (θ, τ) estimation. Since the result in (23) shows asymptotic illustration, therefore we can define the distribution of the TSQRDPD estimator as follows:

$$(NT)^{\frac{1}{2}}(\hat{\omega} - \omega_0) \xrightarrow{d} N(0, \Re\zeta\Re'),$$

where $\Re = Q_{zz}^{-1}\mathbb{N}(\Phi_0)'[I, -Q^* Q_1^{-1}\Gamma_{01}, -Q^* Q_2^{-1}\Gamma_{02}, ..., -Q^* Q_{\mathcal{W}}^{-1}\Gamma_{0\mathcal{W}}]$. We define $\psi_\theta(\mathbb{K}_{it}) = [\psi_\theta(v_{it}, \psi_\theta(V_{1it-1}), \psi_\theta(V_{2it-1}), ..., \psi_\theta(V_{\mathcal{W}it-1}))]'$, where $\mathbb{K}_{it}$ can be defined as the general form of the components for structural error term. $Q_{zz} = \mathbb{N}(\Phi_0)' Q^* \mathbb{N}(\Phi_0)$, hence, $\mathbb{W}_{it} = \psi_\theta(\mathbb{K}_{it}) \otimes x_{it}$.

The term is defined as $\zeta = E(\Delta \otimes x_{it}x'_{it})$, where Δ is considered to be the general form for matrix of the following terms: $\psi_\theta(\mathbb{K}_{rit})\psi_\theta(\mathbb{K}_{r^*it})$, where $(\mathbb{K}_{rit}) = v_{it}$ and $(\mathbb{K}_{r^*it}) = V_{r-1, it-1}$, and based on the models in (12) and (13), r would be $2 \leq r \leq (\mathcal{W} + 1)$. According to Assumption 3 in Kim and Muller [13] along with the variance-covariance matrix derivation performed by Buchinsky [5], we can present our variance-covariance matrix by utilizing plug-in technique as follows:

$$\hat{\zeta} = (NT)^{\frac{1}{2}} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\Delta} x_{it} x'_{it}, \tag{24}$$

where $\hat{\Delta}$ is verified to be the estimator of general form for matrix for the following terms: $\psi_\theta(\hat{\mathbb{K}}_{rit})\psi_\theta(\hat{\mathbb{K}}_{r^*it})$, where $(\hat{\mathbb{K}}_{rit}) = v_{it}$ and $(\hat{\mathbb{K}}_{r^*it}) = V_{r-1, it-1}$, and they follow the models in (1) and (2), where r would be

$2 \leq r \leq (\mathcal{W}+1)$. It should be also mentioned that by utilizing the Assumptions 1-3, $\hat{\Re} \xrightarrow{d} \Re$ and $\hat{\zeta} \xrightarrow{d} \zeta$ which imply that variance-covariance of TSQRDPD structural parameter is consistent.

4. Conclusion

A new estimation procedure for DPD to tackle the issue of endogeneity through TSQRDPD is extensively presented. We obtained the asymptotic distribution for TSQRDPD structural parameter. We also presented the general form of variance-covariance matrix and proved its consistency.

References

[1] T. Amemiya, Two stage absolute deviations estimators, Econometrica 50(3) (1982), 689-711.

[2] D. Andrews, Empirical process methods in econometrics, Handbook of Econometrics, Elsevier Science, New York, Vol. 4, 1994, pp. 2248-2292.

[3] D. H. Autor, S. N. Houseman and S. P. Kerr, The effect of work first job placements on the distribution of earnings: an instrumental variable quantile regression approach, Journal of Labor Economics 35(1) (2017), 149-190.

[4] P. Bickel, One-step Huber estimates in the linear model, J. Amer. Statist. Assoc. 70(350) (1975), 428-434.

[5] M. Buchinsky, Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study, J. Econometrics 68(2) (1995), 303-338.

[6] R. Carroll and D. Ruppert, Robust estimation in heteroscedastic linear models, Ann. Statist. 10(2) (1982), 429-441.

[7] L. Chen and S. Portnoy, Two-stage regression quantiles and two-stage trimmed least squares estimators for structural equation models, Comm. Statist. Theory Methods 25(5) (1996), 1005-1032.

[8] V. Chernozhukov and C. Hansen, Instrumental variable quantile regression: a robust inference approach, J. Econometrics 142(1) (2008), 379-398.

[9] M. Dudely, Central limit theorems for empirical measures, Ann. Probab. 6 (1979), 909-911.

[10] A. Galvao, Quantile regression for dynamic panel data with fixed effects, J. Econometrics 164 (2011), 142-157.

[11] J. Gu and R. Koenker, Unobserved heterogeneity in income dynamics: an empirical Bayes perspective, J. Bus. Econom. Statist. 35(1) (2017), 1-16.

[12] G. Kemp, Least absolute error difference estimation of a single equantion from simultaneous equations system, Unviversity of Essex, Essex, 1999.

[13] T. Kim and C. Muller, Two-stage Quantile Regression, University of Nottingham, Nottingham, 2000.

[14] T. Kim and C. Muller, Inconsistency transmission and variance reduction in two-stage quantile regression, Canad. J. Statist. (2017), in press.

[15] T. Kim and C. Muller, Two-stage Quantile Regression when the First Stage is based on Quantile Regression, Nottingham University, Nottingham, 2003.

[16] T. Kim and C. Muller, Two-stage quantile regression when the first stage is based on quantile regression, Econom. J. 7 (2004), 218-231.

[17] R. Koenker and G. Bassett, Regression quantiles, Econometrica 46(1) (1978), 33-50.

[18] L. Ma and R. Koenker, Quantile regression for recursive structural, J. Econometrics 134 (2005), 471-506.

[19] T. MaCurdy and C. Timmins, Bounding the influence of attrition on intertemporal wage variation in NSLY, University of Stanford, Stanford, 2000.

[20] R. Moon and M. Weidner, Dynamic linear panel regression models with interactive fixed effects, Econom. Theory 33(1) (2017), 158-195.

[21] D. Pollard, Asymptotics via empirical process, Statist. Sci. 4 (1989), 341-366.

[22] D. Pollard, Asymptotics for least absolute deviations regression estimators, Econom. Theory 7 (1991), 186-199.

[23] J. Powell, The asymptotic normality of two-stage least absolute deviations estimators, Econometrica 51(5) (1983), 1569-1575.

[24] A. Vapnik and Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Theory Probab. Appl. 16(2) (1971), 264-280.