

Yan, Karen X.; Li, Qi

**Article**

## Nonparametric estimation of a conditional quantile function in a fixed effects panel data model

Journal of Risk and Financial Management

**Provided in Cooperation with:**

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Yan, Karen X.; Li, Qi (2018) : Nonparametric estimation of a conditional quantile function in a fixed effects panel data model, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 11, Iss. 3, pp. 1-10, <https://doi.org/10.3390/jrfm11030044>

This Version is available at:

<https://hdl.handle.net/10419/238885>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



Article

# Nonparametric Estimation of a Conditional Quantile Function in a Fixed Effects Panel Data Model

Karen X. Yan <sup>1</sup> and Qi Li <sup>1,2,\*</sup>

<sup>1</sup> Department of Economics, Texas A&M University, College Station, TX 77845, USA; yanxueqing@tamu.edu

<sup>2</sup> International School of Economics and Management (ISEM), Capital University of Economics and Business, Beijing 100070, China

\* Correspondence: qi-li@tamu.edu

Received: 11 July 2018; Accepted: 1 August 2018; Published: 3 August 2018



**Abstract:** This paper develops a nonparametric method to estimate a conditional quantile function for a panel data model with an additive individual fixed effects. The proposed method is easy to implement, it does not require numerical optimization and automatically ensures quantile monotonicity by construction. Monte Carlo simulations show that the proposed estimator performs well in finite samples.

**Keywords:** nonparametric method; conditional quantile function; panel data

**JEL Classification:** C14; C21

## 1. Introduction

Using nonparametric techniques to estimate econometric models has received increasing attention among econometricians in recent decades (see, for example, [Pagan and Ullah \(1999\)](#); [Hall et al. \(2007\)](#); [Belloni et al. \(2016\)](#); [Lin et al. \(2015\)](#); [Li et al. \(2013\)](#); [Firpo et al. \(2009\)](#) and [Firpo et al. \(2018\)](#) for the literature of nonparametric methods and applications). The most popular nonparametric model is the conditional mean regression model. However, compared with a conditional mean function, a conditional quantile regression function, when evaluated at different quantiles, can reveal an entire distributional relationship between the covariates and the response variable. Quantile regression therefore has many useful applications in economics and finance. For example, in risk and financial management, researchers are more concerned about the uncertainty or the risk of an asset, which can be characterized by its left tail behavior (corresponding to the lower quantiles) (see [Al Rahahleh and Bhatti \(2017\)](#); [Al Rahahleh et al. \(2017\)](#); [Nguyen and Bhatti \(2015\)](#); [Al Rahahleh et al. \(2016\)](#); [Bartram et al. \(2018\)](#); [Al Shubiri and Jamil \(2018\)](#) for the literature on idiosyncratic risk), and quantile regression can play an important role in this line of research.

The existing work on nonparametric estimation of quantile functions mostly focuses on cross-sectional data, or weakly dependent stationary data processes. Nonparametric estimation of conditional quantile functions with panel data is more difficult when there exists fixed effects term that is correlated with covariates. In this paper, we consider the following nonparametric panel data model with individual fixed effects:

$$Y_{it} = \alpha_i + m(X_{it}) + \epsilon_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T, \quad (1)$$

where  $Y_{it}$  is the outcome variable,  $X_{it}$  is a scalar<sup>1</sup>,  $\alpha_i$  is the individual fixed effect, it has zero mean and is allowed to be correlated with  $X_{it}$  in an unknown correlation form,  $m(\cdot)$  is smooth but otherwise unspecified function, the idiosyncratic error  $\epsilon_{it}$  is *i.i.d* with zero mean and a finite variance. Given that  $\alpha_i + \epsilon_{it}$  has a zero mean, we have from Equation (1) that  $E(Y_{it}) = E[m(X_{it})]$ . Without loss of generality, we assume that  $E[m(X_{it})] = 0$ .<sup>2</sup>

A key attractive feature of panel data for empirical researchers is that it controls for the unobserved heterogeneity. Equation (1) has been discussed in Henderson et al. (2008), with a focus on the nonparametric estimation and testing of the conditional mean function. Our interest lies in estimating the conditional quantile function of  $Y_{it} - \alpha_i = m(X_{it}) + \epsilon_{it}$  given  $X_{it} = x$ . The application of quantile regression to panel data framework has been a challenging task (see, for example, Koenker (2004); Abrevaya and Dahl (2008); Kato et al. (2012); Harding and Lamarche (2014)). The check-function method and inverse-CDF method are the two main methods in quantile regression analysis, with the former most widely used in literature. One main challenge with the check-function method is that the objective criterion function is non-differentiable and therefore numerical optimization is required. This creates a computational burden. Another drawback of the check-function method is the lack of monotonicity, also known as the quantile crossing problem (see Bassett and Koenker (1982) and He (1997)). Researchers often need to impose shape restrictions or use monotone rearrangement to address the quantile crossing problem (Chernozhukov et al. (2010); Qu and Yoon (2015)).

This paper develops a new quantile regression method for the nonparametric panel data Equation (1) in the spirit of Fang et al. (2018)<sup>3</sup>. The new method exploits the location-scale structure of Equation (1). Note that the conditional  $\tau$ -th quantile function of  $Y_{it} - \alpha_i$  given  $X_{it} = x$ , denoted by  $q_\tau(x)$ , takes a particularly simple closed-form structure:

$$q_\tau(x) = m(x) + Q_\epsilon(\tau), \tag{2}$$

for all  $\tau \in (0, 1)$ , where  $Q_\epsilon(\tau)$  is the  $\tau$ -th quantile of  $\epsilon_{it}$ <sup>4</sup>. Thus, if  $\hat{m}(x)$  is the estimator of  $m(x)$ , then  $q_\tau(x)$  can be estimated by

$$\hat{q}_\tau(x) = \hat{m}(x) + \hat{Q}_\epsilon(\tau), \tag{3}$$

where  $\hat{Q}_\epsilon(\tau)$  is the empirical quantile function of the (normalized) regression residuals.

For estimation, we first use the first-difference transformation to get rid of the individual fixed effect  $\alpha_i$  and estimate the unknown function  $m(\cdot)$  by the series method, we then use deconvolution method to back up the distribution of error term  $\{\epsilon_{it}\}$ , therefore the quantile estimator of  $\epsilon$ . Finally, we exploit the location-scale structure of the first-differenced model to derive the quantile estimator of  $Y_{it} - \alpha_i$ , which is given in Equation (3). The deconvolution step closely relates to the papers by Horowitz and Markatou (1996) and Evdokimov (2010) for the application of the deconvolution method to recover the density of panel data error term. Our approach does not require numerical optimization, is computationally easy to implement, and automatically ensures quantile monotonicity by construction. For asymptotic property of the conditional quantile estimator, as long as the series estimator  $\hat{m}(x)$  and  $\hat{Q}_\epsilon(\tau)$  are consistent<sup>5</sup>, the conditional quantile estimator  $\hat{q}_\tau(x)$  is also consistent by

<sup>1</sup> For ease of exposition, we assume  $X_{it}$  is univariate, the extension to multivariate case can be carried over straightforwardly.

<sup>2</sup> This can be achieved by using de-mean data for the dependent variable, i.e., replacing  $Y_{it}$  by  $Y_{it} - (NT)^{-1} \sum_{j=1}^N \sum_{s=1}^T Y_{js}$  in Equation (1). For notational simplicity, we still use  $Y_{it}$  to denote the dependent variable although it is actually the de-mean version of it.

<sup>3</sup> Recently, Fang et al. (2018) proposes a new nonparametric method for estimating a conditional quantile function with cross-sectional data. We refer readers to Fang et al. (2018) for a detailed discussion.

<sup>4</sup> For ease of exposition, we drop the subscript *it* in  $Q_{\epsilon_{it}}(\tau)$  and use  $Q_\epsilon(\tau)$  to denote the  $\tau$ -th quantile of  $\epsilon_{it}$  in general, since  $\epsilon_{it}$  is an *i.i.d.* sequence.

<sup>5</sup> The consistency can be straightforwardly shown using similar arguments as in Fang et al. (2018) and Horowitz and Markatou (1996).

Equation (3) and the continuous mapping theorem. While we do not provide theoretical underpinnings for the proposed quantile estimator, Monte Carlo simulation results show that the estimator performs well in finite samples.

The remainder of the paper is organized as follows. Section 2 gives a detailed description of the methodology. Section 3 presents a Monte Carlo simulation to examine the finite-sample performance of the proposed quantile estimator. Section 4 considers an extension where the error is heteroskedastic. Section 5 concludes the paper.

## 2. Methodology

In this section, we describe the three-step procedure to estimate the conditional quantile function  $q_\tau(x)$ .

**STEP 1.** Use the first-difference to get rid of individual fixed effects and estimate  $m(\cdot)$  by the nonparametric series method.

First differencing Equation (1), we have that

$$Y_{it} - Y_{i,t-1} = m(X_{it}) - m(X_{i,t-1}) + \epsilon_{it} - \epsilon_{i,t-1}, \quad i = 1, \dots, N; t = 2, \dots, T. \quad (4)$$

Note that despite if one uses a de-mean dependent variable or not, it leads to the same first-differenced Equation (4) because any additive constant will be wiped out by first-difference transformation.

Let  $P^K(X_{it}) = [\zeta_1(X_{it}), \zeta_2(X_{it}), \dots, \zeta_K(X_{it})]'$  denote the  $K \times 1$  dimensional basis functions, where  $K$  is the number of basis functions. For example, we may choose power series base function so that  $[\zeta_1(X_{it}), \zeta_2(X_{it}), \dots, \zeta_K(X_{it})]' = [X_{it}, X_{it}^2, \dots, X_{it}^K]'$ , or we can choose spline base function. By the approximation property of series basis function, there exists an  $K \times 1$  vector of constants  $\beta$  such that  $\sup_{x \in S} |P^K(x)' \beta - m(x)| \rightarrow 0$  as  $K \rightarrow \infty$ , where  $S \in \mathcal{R}$  is a compact support of  $X_{it}$ . In practice, one can estimate  $\beta$  by the least squares method based on

$$\begin{aligned} \Delta Y_{it} &= Y_{it} - Y_{i,t-1} \\ &= [P^K(X_{it})' - P^K(X_{i,t-1})'] \beta + v_{it} \\ &= \Delta P_{it} + v_{it}, \end{aligned} \quad (5)$$

where  $v_{it} = \epsilon_{it} - \epsilon_{i,t-1} + m(X_{it}) - m(X_{i,t-1}) - [P^K(X_{it})' - P^K(X_{i,t-1})'] \beta$ ,  $\Delta Y_{it} \equiv Y_{it} - Y_{i,t-1}$ ,  $\Delta P_{it} \equiv P^K(X_{it}) - P^K(X_{i,t-1})$ , and  $v_{it} = \epsilon_{it} - \epsilon_{i,t-1} + m(X_{it}) - m(X_{i,t-1}) - [P^K(X_{it})' - P^K(X_{i,t-1})'] \beta$ .

We estimate  $\beta$  by applying the OLS to Equation (5), yielding that

$$\hat{\beta} = (\Delta P' \Delta P)^{-1} \Delta P' \Delta Y,$$

where  $\Delta P = [\Delta P_{2'} \dots \Delta P_{N'}]'$  is an  $N(T-1) \times K$  matrix of base functions,  $\Delta P_i = [\Delta P_{i2} \dots \Delta P_{iT}]'$  is an  $(T-1) \times K$  matrix,  $\Delta Y = [\Delta Y_{2'} \dots \Delta Y_{N'}]'$  is an  $N(T-1) \times 1$  vector of outcome variables, and  $\Delta Y_i = [\Delta Y_{i2} \dots \Delta Y_{iT}]'$  is an  $(T-1) \times 1$  vector.

We therefore obtain the series estimator of  $m(X_{it})$ :

$$\hat{m}(X_{it}) = P^K(X_{it})' \hat{\beta}, \quad i = 1, \dots, N; t = 1, \dots, T.$$

**STEP 2.** Let  $f_\epsilon(\cdot)$  denote the density of  $\epsilon_{it}$ . In this step, we use the deconvolution method to recover  $f_\epsilon(\cdot)$ .

From Step 1, one can obtain the estimator of  $u_{it} = \epsilon_{it} - \epsilon_{i,t-1}$  by  $\hat{u}_{it} \equiv Y_{it} - Y_{i,t-1} - \hat{m}(x_{it}) - \hat{m}(x_{i,t-1}) = \hat{\epsilon}_{it} - \hat{\epsilon}_{i,t-1}$ .

To see how the density of  $\epsilon_{it}$  can be estimated, let  $\phi_u(t) = E(e^{tu_{it}})$  and  $\phi_\epsilon = E(e^{t\epsilon_{it}})$  denote the characteristic functions of  $u_{it}$  and  $\epsilon_{it}$ , respectively, where  $\iota = \sqrt{-1}$ . Assume that the distribution of  $\epsilon_{it}$  is such that  $\phi_\epsilon$  is real and positive for all  $t \in \mathcal{R}$ . Then, it is easy to see that

$$\begin{aligned} \phi_u(t) &= E(e^{tu_{it}}) \\ &= E(e^{t\epsilon_{it} - t\epsilon_{i,t-1}}) \\ &= E(e^{t\epsilon_{it}})E(e^{-t\epsilon_{i,t-1}}) \\ &= E(e^{t\epsilon_{it}})E(e^{t\epsilon_{i,t-1}}) \\ &= [E(e^{t\epsilon_{it}})]^2 \\ &\equiv [\phi_\epsilon(t)]^2, \end{aligned}$$

where in the third equality we use the independence of  $\epsilon_{it}$  and  $\epsilon_{i,t-1}$ , and the fourth equality uses the symmetry of  $\eta_{i,t-1}$ .

Therefore,

$$\phi_\epsilon(t) = \sqrt{\phi_u(t)}. \tag{6}$$

We propose the following steps to obtain the density estimate of  $\epsilon_{it}$ :

- (1) Estimate  $\phi_u(t)$  by

$$\hat{\phi}_u(t) = \frac{1}{NT} \sum_{i=1}^N \sum_{s=1}^T e^{t\eta_{is}}. \tag{7}$$

- (2) By Equations (6) and (7), we estimate  $\phi_\epsilon(t)$  by

$$\hat{\phi}_\epsilon(t) = \sqrt{\hat{\phi}_u(t)}.$$

- (3) By the deconvolution method, we estimate  $f_\epsilon(\cdot)$  by

$$\hat{f}_\epsilon(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \Phi_k\left(\frac{t}{T_n}\right) \hat{\phi}_\epsilon(t) dt, \quad z \in \mathcal{R}, \tag{8}$$

where  $\Phi_k\left(\frac{t}{T_n}\right)$  is the Fourier transform of the kernel function  $k(x) = \frac{\sin\pi x}{\pi x}$  with bandwidth  $\frac{1}{T_n}$ , and

$$\Phi_k(t) = \begin{cases} 0, & |t| \geq \frac{1}{2}, \\ \frac{1}{2}, & |t| = \frac{1}{2}, \\ 1, & |t| \leq \frac{1}{2}. \end{cases}$$

**STEP 3.** We estimate  $Q_\epsilon(\tau)$  by  $\hat{Q}_\epsilon(\tau)$  such that for  $\tau \in (0, 1)$ ,  $\hat{Q}_\epsilon(\tau)$  satisfies the following condition:

$$\tau = \int_{-\infty}^{\hat{Q}_\epsilon(\tau)} \hat{f}_\epsilon(x) dx.$$

Therefore, for  $\tau \in (0, 1)$ , the  $\tau$ -th conditional quantile estimator of  $Y_{it} - \alpha_i$ , given  $X_{it} = x$ , is estimated by

$$\hat{q}_\tau(x) = \hat{m}(x) + \hat{Q}_\epsilon(\tau),$$

where  $\hat{m}(x)$  and  $\hat{Q}_\epsilon(\tau)$  are estimated in Steps 1 and 2, respectively.

**Remark 1.** In Step 1, the consistency estimation of  $m(x)$  requires that as  $NT \rightarrow \infty$ ,  $K \rightarrow \infty$  and  $K/(NT) \rightarrow 0$ . In series estimation,  $K/(NT)$  plays a role similar to the bandwidth  $h$  in kernel methods. In practice, one can use Mallows's  $C_L$  or leave-one-out cross-validation method to determine the series term  $K$ . We refer readers to [Li and Racine \(2007\)](#) for details.

**Remark 2.** Note that, in Step 2, assuming  $\phi_\epsilon(t)$  is real and is equivalent to assuming that the density of  $\epsilon_{it}$  is symmetric around 0. We are using the assumption that  $\phi_{\epsilon_{it}}$  is positive in deriving Equation (6).

**Remark 3.** In Step 2, the smoothing parameter  $T_n$  depends on the sample size  $n = NT$ . To guarantee that  $\hat{\phi}_\epsilon(t)$  uniformly converges to  $\phi_\epsilon(t)$  over  $[-T_n, T_n]$  at a geometric rate with respect to the sample size  $n$ , [Hu and Ridder \(2010\)](#) suggests that we can choose  $T_n$  such that

$$T_n = c \left( \frac{n}{\log(n)} \right)^\gamma, \quad \gamma \in \left( 0, \frac{1}{2} \right),$$

where  $c > 0$  is a constant.

**Remark 4.** For inference, we recommend using a residual bootstrap method similar to [Fang et al. \(2018\)](#). We leave the proof of validity of such a bootstrap procedure to a future research topic.

### 3. Monte Carlo Simulation

In this section, we conduct Monte Carlo simulations to assess the performance of the proposed conditional quantile estimator.

We consider the following data generating process (DGP):

$$Y_{it} = \alpha_i + 2\sin(X_{it}) + X_{it} + \epsilon_{it},$$

where  $X_{it} = 0.4\alpha_i + \xi_{it}$ , where  $\xi_{it}$  is *i.i.d.*  $\sim \text{uniform}[-1, 1]$ ,  $\alpha_i$  is *i.i.d.*  $\sim \text{uniform}[-1, 1]$ . We consider two distributions for  $\epsilon_{it}$ : (i)  $\epsilon_{it}$  is *i.i.d.*  $\sim N(0, 1)$ ; (ii)  $\epsilon_{it}$  is *i.i.d.*  $\sim t(3)$  (a *t*-distribution with degree of freedom 3).

We conduct 2000 Monte Carlo replications for samples of size  $N = 100, 200, 400$  with  $T = 10$ . We report mean squared error (MSE) of three estimators: (1) the series estimator  $\hat{m}(x)$  with  $MSE(\hat{m}) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (\hat{m}(X_{it}) - m(X_{it}))^2$ , (2) the quantile estimator  $\hat{Q}_\tau(\epsilon)$  with  $MSE(\hat{Q}_\epsilon(\tau)) = [\hat{Q}_\epsilon(\tau) - Q_\epsilon(\tau)]^2$ , and (3) the conditional quantile estimator  $\hat{q}_\tau(x)$  with  $MSE(\hat{q}_\tau) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (\hat{q}_\tau(X_{it}) - q_\tau(X_{it}))^2$ . For each of the three quantities above, we average them over the 2000 replications.

We first examine the performance of the deconvolution method for recovering the density of error terms. As an illustration, we only present the result ([Figure 1](#)) for the case of  $\epsilon_{it} \sim N(0, 1)$ , with sample size  $N = 100, T = 10$ . We examine the sensitivity of the estimated density to the choice of different bandwidths. We set  $c = 1$ , and  $\gamma = \frac{1}{8}, \frac{3}{16}, \frac{1}{4}, \frac{3}{8}$ . It can be seen from [Figure 1](#) that the performance of the deconvolution method can be somewhat sensitive to the choice of bandwidth. This is a well known problem of the deconvolution method, not a particular problem to our approach. When  $\gamma$  is small, say  $\gamma = \frac{1}{8}$ , the estimated density is flatter than the true density. However, generally, the estimated density tracks the true density<sup>6</sup>.

[Tables 1](#) and [2](#) report the Mean MSE of  $\hat{m}$ ,  $\hat{Q}_\epsilon(\tau)$  and  $\hat{q}_\tau$ . It can be seen that, as sample size doubles, MSEs of  $\hat{m}$ ,  $\hat{Q}_\epsilon(\tau)$ , and  $\hat{q}_\tau$  decrease by about  $\frac{1}{2}$ , which indicates that the proposed estimator behaves well.

---

<sup>6</sup> There is no rule-of-thumb to choose the optimal bandwidth in the deconvolution method. In practice, researchers can try different bandwidths as a robust check to see how results vary across the different bandwidths.

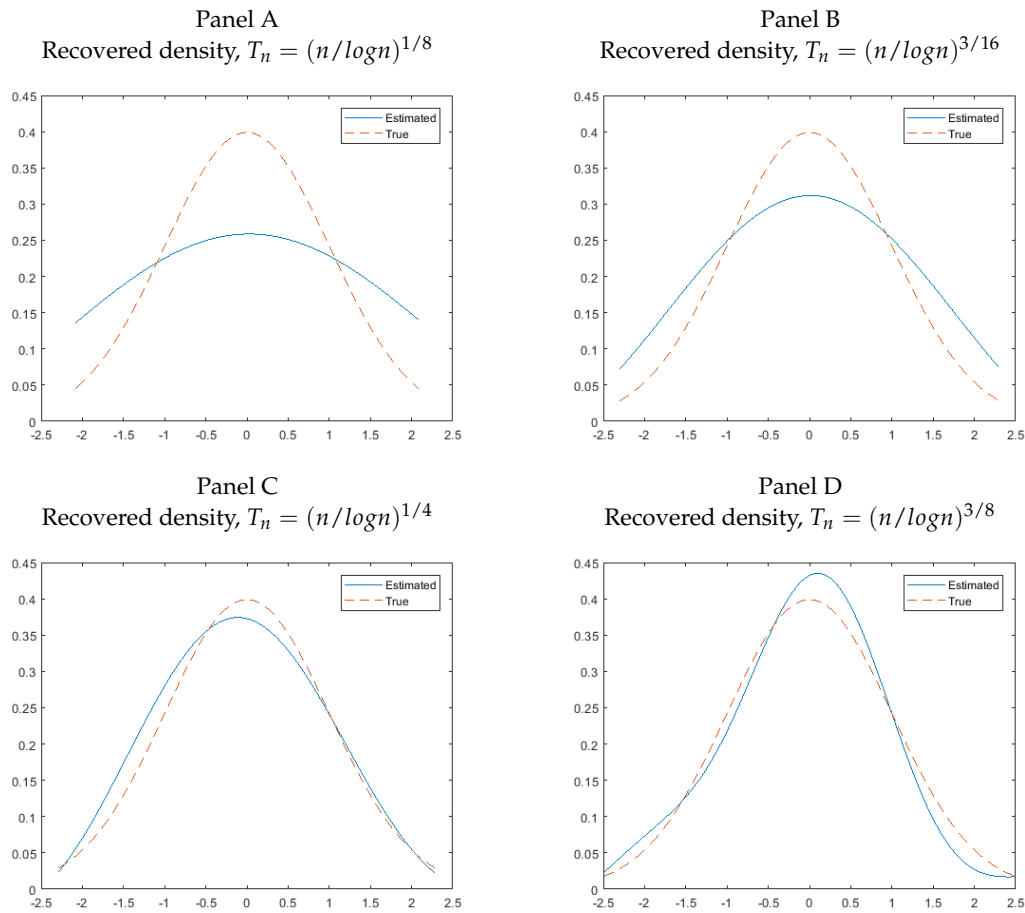


Figure 1. Recovered densities across different bandwidths and homoskedastic symmetric normal errors.

Table 1. Mean MSE ( $\times 100$ ),  $N(0, 1)$  Errors.

Sample Size ( $N, T$ )	Estimators						
	$MSE(\hat{m})$	$MSE(\hat{Q}_\epsilon(\tau))$			$MSE(\hat{q}_\tau)$		
		$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$
(100, 10)	0.0149	0.0037	0.0022	0.0006	0.0204	0.0185	0.0163
(200, 10)	0.0092	0.0012	0.0007	0.0002	0.0107	0.0102	0.0096
(400, 10)	0.0048	0.00051	0.00028	0.000082	0.0052	0.0050	0.0048

Table 2. Mean MSE ( $\times 100$ ),  $t(3)$  Errors.

Sample Size ( $N, T$ )	Estimators						
	$MSE(\hat{m})$	$MSE(\hat{Q}_\epsilon(\tau))$			$MSE(\hat{q}_\tau)$		
		$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.4$
(100, 10)	0.0139	0.0423	0.0235	0.0065	0.0642	0.0433	0.0235
(200, 10)	0.0094	0.0210	0.0128	0.0036	0.0304	0.0222	0.0130
(400, 10)	0.0048	0.0091	0.0063	0.0019	0.0104	0.0112	0.0067

#### 4. Extension: Conditional Heteroskedastic Error Case

In this section, we consider an extension where the error term is conditional heteroskedastic. Specifically, we generalize Equation (1) to the following case<sup>7</sup>:

$$Y_{it} = \alpha_i + m(X_{it}) + \sigma(X_{it})\eta_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T, \tag{9}$$

where  $\sigma(X_{it}) > 0$  is an unknown function,  $\eta_{it}$  is assumed to be *i.i.d* with zero mean, unit variance and independent of  $\{X_{js}\}_{j=1, \dots, N; s=1, \dots, T}$ . Without loss of generality, we assume that  $E[m(X_{it})] = 0$  (similar to the conditional homoskedasticity case).

Define  $\epsilon_{it} \equiv \sigma(X_{it})\eta_{it}$ . The conditional  $\tau$ -th quantile function of  $Y_{it} - \alpha_i$  given  $X_{it} = x$ , denoted by  $q_\tau(x)$ , takes the following closed-form structure:

$$q_\tau(x) = m(x) + Q_{\epsilon|X=x}(\tau) = m(x) + \sigma(x)Q_\eta(\tau), \tag{10}$$

for all  $\tau \in (0, 1)$ , where  $Q_{\epsilon|X=x}(\tau) = \sigma(x)Q_\eta(\tau)$ , and  $Q_\eta(\tau)$  is the (un-conditional)  $\tau$ -th quantile of  $\eta_{it}$ .

**Remark 5.** In deriving Equation (10), we use the fact that  $Q_{\epsilon|X=x}(\tau) = \sigma(x)Q_\eta(\tau)$  because  $\sigma(x) > 0$  and  $X_{it}$  and  $\eta_{it}$  are independent with each other.

**Remark 6.** Noting that, due to the independence between  $X_{it}$  and  $\epsilon_{js}$ , we have that  $\epsilon_{it}|_{X_{it}=x} = \epsilon_{it}|_{X_{it}=X_{i,t-1}=x} = \sigma(x)\eta_{it}$ , and this implies that  $f_{\epsilon_{it}|X_{it}, X_{i,t-1}}(u) = f_{\epsilon_{it}|X_{it}}(u)$  (conditional independence property).

We propose the following three-step procedure to estimate the conditional quantile function of  $Y_{it} - \alpha_i = m(X_{it}) + \sigma(X_{it})\eta_{it}$  given  $X_{it} = x$ .

**STEP 1.** We obtain  $\hat{m}(X_{it}) = P^K(X_{it})'\hat{\beta}$  by exactly the same procedure as in Step 1 of the conditional homoskedastic error case.

**STEP 2.** We use the deconvolution method to estimate  $f_{\epsilon_{it}|X_{it}=x}(\cdot)$ , the conditional density of  $\epsilon_{it}$  given  $X_{it} = x$ . Define  $\Delta Y_{it} \equiv Y_{it} - Y_{it-1}$ . Assuming that the density of  $\eta_{it}$  is symmetric around zero,<sup>8</sup> and note that  $[m(X_{it}) - m(X_{i,t-1})]_{X_{it}=x, X_{i,t-1}=x} = m(x) - m(x) = 0$ , we have

$$\begin{aligned} \phi_{\Delta Y_{it}}(s|x) &\equiv E[\exp(\iota s \Delta Y_{it})|X_{it} = X_{i,t-1} = x] \\ &= E[\exp[\iota s(\epsilon_{it} - \epsilon_{i,t-1})|X_{it} = X_{i,t-1} = x] \\ &= \left\{ E[\exp(\iota s \epsilon_{it})|X_{it} = x] \right\} \left\{ E[\exp(-\iota s \epsilon_{i,t-1})|X_{i,t-1} = x] \right\} \\ &= \left\{ E[\exp(\iota s \epsilon_{it})|X_{it} = x] \right\} \left\{ E[\exp(\iota s \epsilon_{i,t-1})|X_{i,t-1} = x] \right\} \\ &= \left\{ E[\exp(\iota s \epsilon_{it})|X_{it} = x] \right\}^2 \\ &\equiv \phi_{\epsilon_{it}}^2(s|x), \end{aligned} \tag{11}$$

where  $\iota = \sqrt{-1}$ , the third equality uses the conditional independence property as described in Remark 6, and in the fourth equality we use the symmetry of  $\epsilon_{i,t-1}|_{X_{i,t-1} = x} = \sigma(x)\eta_{i,t-1}$ , and  $\eta_{i,t-1}$  is symmetric around zero.

<sup>7</sup> Fang et al. (2018) also considers the same form of heteroskedastic error as described here.

<sup>8</sup> This implies the conditional density of  $\epsilon_{it}$  given  $X_{it} = x$  is symmetric, since, given that  $\epsilon_{it}|_{X_{it}=x} = \sigma(x)\eta_{it}$ , the symmetry of  $\eta_{it}$  is equivalent to the symmetry of  $\epsilon_{it}$ .



Under the assumption that  $\phi_\epsilon(s|x)$  is positive, the above equation implies that  $\phi_{\epsilon_{it}}(s|x) = \sqrt{\phi_{\Delta Y_{it}}(s|x)}$ . The left-hand side of Equation (11) can be estimated from data:

$$\hat{\phi}_{\Delta Y_{it}}(s|x) = \frac{\sum_{i=1}^N \sum_{t=2}^T \exp[ts(Y_{it} - Y_{i,t-1})] K\left(\frac{X_{it}-x}{h_1}\right) K\left(\frac{X_{i,t-1}-x}{h_2}\right)}{\sum_{i=1}^N \sum_{t=2}^T K\left(\frac{X_{it}-x}{h_1}\right) K\left(\frac{X_{i,t-1}-x}{h_2}\right)}.$$

Therefore, we estimate  $\phi_{\epsilon_{it}}(s|x)$  by  $\hat{\phi}_{\epsilon_{it}}(s|x) = \sqrt{\hat{\phi}_{\Delta Y_{it}}(s|x)}$ . Let  $f_{\epsilon_{it}|X_{it}=x}(\cdot)$  denote the conditional density of  $\epsilon_{it} = \sigma(X_{it})\eta_{it}$  given  $X_{it} = x$ . Then, using the deconvolution method as in the homoskedastic case, one can recover  $\hat{f}_{\epsilon_{it}|X_{it}=x}(\cdot)$  using  $\hat{\phi}_{\epsilon_{it}}(s|x)$  as in Equation (8). We use  $\hat{f}_{\epsilon_{it}|X_{it}=x}(\cdot)$  to denote the resulting estimator of  $f_{\epsilon_{it}|X_{it}=x}(\cdot)$ .

**STEP 3.** Let  $\hat{Q}_{\epsilon|X=x}(\tau)$  denote the estimate of  $Q_{\epsilon|X=x}(\tau)$ ,  $\tau \in (0, 1)$ . The following identity

$$\tau = \int_{-\infty}^{Q_{\epsilon|X=x}(\tau)} f_{\epsilon|X=x}(z) dz$$

suggests that we can obtain  $\hat{Q}_{\epsilon|X=x}(\tau)$  based on the following equation:

$$\tau = \int_{-\infty}^{\hat{Q}_{\epsilon|X=x}(\tau)} \hat{f}_{\epsilon|X=x}(z) dz,$$

where  $\hat{f}_{\epsilon|X=x}(z)$  is estimated from Step 2.

By Equation (10), the  $\tau$ -th conditional quantile estimator of  $Y_{it} - \alpha_i$ , given  $X_{it} = x$ , is estimated by

$$\hat{q}_\tau(x) = \hat{m}(x) + \hat{Q}_{\epsilon|X=x}(\tau), \quad \tau \in (0, 1),$$

where  $\hat{m}(x) = P^K(x)' \hat{\beta}$  is obtained in Step 1, and  $\hat{Q}_{\epsilon|X=x}(\tau)$  is obtained in Step 3.

**Remark 7.** Note that, in the last step, we estimate the  $\tau$ -th quantile of  $\epsilon_{it} \equiv \sigma(X_{it})\eta_{it}$  directly, instead of estimating the unknown function  $\sigma(\cdot)$  and  $Q_\eta(\tau)$  separately (e.g., Fang et al. (2018)).

## 5. Conclusions

In this paper, we propose an easy-to-implement nonparametric method to estimate conditional quantile functions in a fixed effects panel data model. There are many directions that one can extend the results of this paper to more general settings. For example, one can allow for panel non-stationary data as considered in Chen and Khan (2008) or allow for the covariate  $X_{it}$  to be endogenous. We leave these as possible future research topics.

**Author Contributions:** The two authors both contribute to the project formulation and paper preparation.

**Funding:** This research received no external funding.

**Acknowledgments:** We thank three anonymous referees for helpful and constructive comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Abrevaya, Jason, and Christian M. Dahl. 2008. The effects of birth inputs on birthweight. *Journal of Business & Economic Statistics* 26: 379–97.
- Al Rahahleh, Naseem, Iman Adeinat, and M. Ishaq Bhatti. 2016. On ethnicity of idiosyncratic risk and stock returns puzzle. *Humanomics* 32: 48–68. [CrossRef]
- Al Rahahleh, Naseem, and M. Ishaq Bhatti. 2017. Co-movement measure of information transmission on international equity markets. *Physica A: Statistical Mechanics and its Applications* 470: 119–31. [CrossRef]

- Al Rahahleh, Naseem, M. Ishaq Bhatti, and Iman Adeinat. 2017. Tail dependence and information flow: Evidence from international equity markets. *Physica A: Statistical Mechanics and its Applications* 474: 319–29. [CrossRef]
- Al Shubiri, Faris Nasif, and Syed Ashsan Jamil. 2018. The impact of idiosyncratic risk of banking sector on oil, stock market, and fiscal indicators of Sultanate of Oman. *International Journal of Engineering Business Management* 10: doi:10.1177/1847979017749043. [CrossRef]
- Bartram, Söhnke M., Gregory W. Brown, and René M. Stulz. 2018. *Why Has Idiosyncratic Risk Been Historically Low in Recent Years?* National Bureau of Economic Research Working Paper NO. 24270, Cambridge, MA, USA: National Bureau of Economic Research, January.
- Bassett, Gilbert, Jr., and Roger Koenker. 1982. An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association* 77: 407–15. [CrossRef]
- Belloni, Alexandre, Victor Chernozhukov, and Ivan Fernández-Val. 2016. Conditional quantile processes based on series and many regressors. *arXiv*, arXiv:1105.6154.
- Chen, Songnian, and Shakeeb Khan. 2008. Semiparametric estimation of nonstationary censored panel data models with time varying factor loads. *Econometric Theory* 24: 1149–73. [CrossRef]
- Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon. 2010. Quantile and probability curves without crossing. *Econometrica* 78: 1093–25.
- Evdokimov, Kirill. 2010. *Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity*. Working Paper, Princeton, NJ, USA: Princeton University.
- Fang, Zheng, Qi Li, and Karen Yan. 2018. *A Simple Nonparametric Method for Estimation and Inference of Conditional Quantile Functions*. Working Paper. Available online: <https://ssrn.com/abstract=3223015> (accessed on 3 August 2018).
- Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica* 77: 953–73.
- Firpo, Sergio P., Nicole M. Fortin, and Thomas Lemieux. 2018. Decomposing wage distributions using recentered influence function regressions. *Econometrics* 6: 28. [CrossRef]
- Hall, Peter, Qi Li, and Jeffrey S. Racine. 2007. Nonparametric estimation of regression functions in the presence of irrelevant regressors. *The Review of Economics and Statistics* 89: 784–89. [CrossRef]
- Harding, Matthew, and Carlos Lamarche. 2014. Estimating and testing a quantile regression model with interactive effects. *Journal of Econometrics* 178: 101–13. [CrossRef]
- He, Xuming. 1997. Quantile curves without crossing. *The American Statistician* 51: 186–92.
- Henderson, Daniel J., Raymond J. Carroll, and Qi Li. 2008. Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics* 144: 257–75. [CrossRef] [PubMed]
- Horowitz, Joel L., and Marianthi Markatou. 1996. Semiparametric estimation of regression models for panel data. *The Review of Economic Studies* 63: 145–68. [CrossRef]
- Hu, Yingyao, and Geert Ridder. 2010. On deconvolution as a first stage nonparametric estimator. *Econometric Reviews* 29: 365–96. [CrossRef]
- Kato, Kengo, Antonio F. Galvao, and Gabriel V. Montes-Rojas. 2012. Asymptotics for panel quantile regression models with individual effects. *Journal of Econometrics* 170: 76–91. [CrossRef]
- Koenker, Roger. 2004. Quantile regression for longitudinal data. *Journal of Multivariate Analysis* 91: 74–89. [CrossRef]
- Li, Qi, Juan Lin, and Jeffrey S. Racine. 2013. Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business & Economic Statistics* 31: 57–65.
- Li, Qi, and Jeffrey Scott Racine. 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton: Princeton University Press.
- Lin, Wei, Zongwu Cai, Zheng Li, and Li Su. 2015. Optimal smoothing in nonparametric conditional quantile derivative function estimation. *Journal of Econometrics* 188: 502–13. [CrossRef]
- Nguyen, Cuong, and M. Ishaq Bhatti. 2015. Investor sentiment and idiosyncratic volatility puzzle: Evidence from the chinese stock market. *Journal of Stock and Forex Trading* 4: 2.

- Pagan, Adrian, and Aman Ullah. 1999. *Nonparametric Econometrics*. Cambridge: Cambridge University Press.
- Qu, Zhongjun, and Jungmo Yoon. 2015. Nonparametric estimation and inference on conditional quantile processes. *Journal of Econometrics* 185: 1–19. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).