Journal of Econometrics 206 (2018) 531–553

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A semiparametric quantile panel data model with an application to estimating the growth effect of FDI[☆]

Zongwu Cai[a,b], Linna Chen[b], Ying Fang[b,*]

[a] Department of Economics, University of Kansas, Lawrence, KS 66045, USA

[b] Wang Yanan Institute for Studies in Economics and Fujian Key Laboratory of Statistical Science, Xiamen University, Xiamen, Fujian 361005, China

ARTICLE INFO

Article history:
Available online 9 July 2018

JEL classification:
C14
C31

Keywords:
Correlated random effect
Foreign direct investment
Panel data
Quantile regression model
Local quasi-likelihood
Semiparametric model
Varying coefficient model

ABSTRACT

This paper estimates the impact of foreign direct investment on economic growth by proposing a new semiparametric quantile panel data model with correlated random effects, in which some of the coefficients are allowed to depend on some smooth economic variables while other coefficients remain constant. A three-stage estimation procedure is proposed to estimate both constant and functional coefficients and their asymptotic properties are investigated. A simple and easily implemented procedure for making inferences is proposed. Monte Carlo simulation is conducted to examine the finite sample performance of the proposed estimators. Finally, using the cross-country panel data, we find a strong empirical evidence of the existence of the absorptive capacity hypothesis, together with another new finding that FDI has much stronger growth effects for countries with fast economic growth than for those with slow economic growth.

1. Introduction

It is well documented in the growth literature that foreign direct investment (FDI) plays an important role in the economic growth process in host countries since FDI is often considered as a vehicle to transfer new ideas, advanced capitals, superior technology and know-how from developed countries to developing countries and so on. However, the existing empirical studies provide contradictory results on whether or not FDI promotes an economic development in host countries.[1] The recent studies in the literature concluded that the mixed empirical evidences may be due to nonlinearities in FDI effects on economic growth and the heterogeneity across countries.

Indeed, it is well recognized by many economists in empirical studies that a standard linear growth model may be inappropriate for investigating the nonlinear effect of FDI on economic development. The nonlinearity in FDI effects is

☆ The authors thank the editor and two anonymous referees for their helpful and constructive comments. Cai's research is partially supported by the National Natural Science Foundation of China grant #71131008 and #71631004 (Key Projects). Fang's research is partially supported by the National Science Fund for Distinguished Young Scholars #71625001.

* Corresponding author.
E-mail addresses: caiz@ku.edu (Z. Cai), linnnachen918@gmail.com (L. Chen), yifst1@xmu.edu.cn (Y. Fang).

1 For example, Blomstrom and Persson (1983), Blomstrom et al. (1992), De Gregorio (1992), Borensztein et al. (1998), De Mello (1999), Ghosh and Wang (2009), Kottaridi and Stengos (2010) among others found positive effects of FDI on promoting the economic growth in various environments. On the other hand, many studies including Haddad and Harrison (1993), Aitken and Harrison (1999), Lipsey (2003), and Carkovic and Levine (2005) failed to find beneficial effects of FDI on the economic growth in host countries. Grog and Strobl (2001) did a meta analysis of 21 studies using the data from 1974 to 2001 that worked on estimating FDI effects on productivity in host countries, of which 13 studies reported positive results, 4 studies reported negative effects and the remaining reported inconclusive evidence.

https://doi.org/10.1016/j.jeconom.2018.06.013
0304-4076/

mainly due to the fact, the so called absorptive capacity in host countries, that host countries need some minimum conditions to absorb the spillovers from FDI.[2] Most existing literature to deal with the nonlinearity issue used simply some parametric nonlinear models, for example, including an interacted term in the regression or running a threshold regression. A parametric nonlinear model has the risk of encountering the model misspecification problem. Misspecified models can lead to biased estimation and misleading empirical results. Recently, Henderson et al. (2011) and Kottaridi and Stengos (2010) adopted nonparametric regression techniques into a growth model. However, due to the curse of dimensionality in a pure nonparametric estimation, such applications are either restricted by the sample size problem or rely heavily on the variable selection which is not an easy task.

The heterogeneity among countries is another concern in cross-country studies. Grog and Strobl (2001) found that whether a cross sectional or time series data model had been used matters for estimating the effect of FDI on the economic growth, because both the cross sectional and time series models cannot control the country-specific heterogeneity. Recent literature focused on using panel data to estimate growth models, which can control the country-specific unobserved heterogeneity using individual effects. However, including individual effects which only allows a location shift for each country, does not have an adequate ability to deal with the heterogeneity effect of FDI on the economic growth across countries. For example, some studies found that empirical results vary depending on whether developed countries are included in the sample. The existing literature to handle this issue is to split sample into groups.[3] Generally speaking, splitting sample can lead to potential theoretical and empirical problems. First, regressing on the split samples separately may lose other parts of information and degrees of freedom, which may lead to inefficient estimation. Secondly, the applied researchers often split sample without following the theoretical guideline on how to select thresholds.

To deal with the aforementioned two issues (nonlinearities and heterogeneity) in a simultaneous fashion, we propose a partially varying-coefficient quantile panel data model with correlated random effects for fixed T to estimate the nonlinear effect of FDI on the economic growth with heterogeneity. Different from the existing literature, we resolve the nonlinearity issue by employing a partially varying-coefficient model which allows some of coefficients to be constant but others, reflecting the effects of FDI on the economic growth, to depend on the country's initial condition. Compared to a fully nonparametric estimation, our model setup can achieve the dimension reduction and accommodate the well recognized economic theory such as the absorptive capacity. In addition to using panel data with individual effects which allows for location shifts for individual countries, we propose a semiparametric conditional quantile regression model instead of commonly used conditional mean models. A conditional quantile model can provide more flexible structures than conditional mean models to characterize heterogeneity among countries. For example, besides including individual effects allowing country-specific heterogeneity, a conditional quantile model allows different growth equations for different quantiles. Therefore, we can take the advantage of utilizing all sample information to identify the effect of FDI on the economic growth without splitting sample according to development stages. Moreover, estimating all quantiles can provide a whole picture of the conditional distribution and avoid the possibly misleading conditional mean results due to the heteroscedasticity in error terms. In other words, using the quantile approach can characterize the different roles of FDI in economic growth for different types of countries.

The application of conditional quantile model to analyze economic and financial data has a long history that can be traced to the seminal papers by Koenker and Bassett (1978, 1982); see the book by Koenker (2005) for more details. Recently, many studies have focused on nonparametric or semiparametric quantile regression models for either independently identically distributed (iid) data or time series data.[4] However, due to the fact that the approach of taking a difference, which is commonly used in conditional mean panel data (linear) models to eliminate individual effects, is invalid in quantile regression settings, even for linear quantile regression model, the literature on quantile panel data models is relatively small. To the best of our knowledge, the paper by Koenker (2004) is the first paper to consider a linear quantile panel data model with fixed effects, where the fixed effects are assumed to have pure location shift effects on the conditional quantiles of the dependent variable but the effects of regressors are allowed to be dependent on quantiles. Koenker (2004) proposed two methods to estimate such a panel data model with fixed effects by assuming that T goes to infinity. The first method is to solve a piecewise linear quantile loss function by using interior point methods and the second one is the penalized quantile regression method, in which the quantile loss function is minimized by adding L_1 penalty on fixed effects. Recently, in a penalized quantile panel data regression model as in Koenker (2004), Lamarche (2010) discussed how to select the tuning parameter, which can control the degree of shrinkage for fixed effects, whereas Galvao (2011) extended the quantile

2 Nunnenkamp (2004) emphasized the importance of the initial condition for host countries to absorb the positive impacts of FDI, Borensztein et al. (1998) found that a threshold stock of human capital in host countries is necessary for them to absorb beneficial effects of advanced technologies brought from FDI, and Hermes and Lensink (2003), Alfaro et al. (2004) and Durham (2004) addressed the local financial market conditions of a country's absorptive capacity.

3 For example, Luiz and De Mello (1999) considered OECD and non-OECD samples and Kottaridi and Stengos (2010) split the whole sample into high-income and middle-income groups.

4 For example, Chaudhuri (1991) studied nonparametric quantile estimation and derived its local Bahadur representation, He et al. (1998), He and Ng (1999), and He and Portnoy (2000) considered nonparametric estimation using splines, De Gooijer and Zerom (2003), Yu and Lu (2004), and Horowitz and Lee (2005) focused on additive quantile models, and Honda (2004) and Cai and Xu (2008) studied varying-coefficient quantile models for time series data. In particular, semiparametric quantile models have attracted increasing research interests during the recent years due to their flexibility. For example, He and Liang (2000) investigated the quantile regression of a partially linear errors-in-variable model, Lee (2003) discussed the efficient estimation of a partially linear quantile regression, and Cai and Xiao (2012) proposed a partially varying-coefficient dynamic quantile regression model, among others.

regression to a dynamic panel data model with fixed effects by employing the lagged dependent variables as instrumental variables and by extending Koenker (2004)'s first method to Chernozhukov and Hansen (2006)'s quantile instrumental variable framework. Finally, Canay (2011) proposed a simple two-stage method to estimate a quantile panel data model with fixed effects. However, the consistency of the estimator in Canay (2011) relies on the assumption of T going to infinity and the existence of an initial $\sqrt{NT}$-consistent estimator in the conditional mean model.

An alternative way to deal with individual effects in a panel data model when T is fixed is to treat them as correlated random effects initiated by Chamberlain (1982, 1984) for the mean regression model. Under the framework of Chamberlain (1982, 1984), to estimate the effect of birth inputs on birth weight, Abrevaya and Dahl (2008) employed a linear quantile panel data model with correlated random effects which are viewed as a linear projection onto some covariates plus an error term. The identification of the effects of covariates only requires two-period information. Furthermore, Gamper-Rabindran et al. (2010) estimated the impact of piped water provision on infant mortality by adopting a linear quantile panel data model with random effects where the random effects were allowed to be correlated with covariates nonparametrically. The model can be estimated through a two-step procedure, in which some conditional quantiles were estimated nonparametrically in the first step and in the second step, the coefficients are estimated by regressing the differenced estimated quantiles on the differenced covariates.

The motivation of this study is to examine the role of FDI in the economic growth process based on the cross-country data from 1970 to 1999 by using the proposed partially varying-coefficient quantile regression model for panel data with correlated random effects. Indeed, this model includes the models in Lee (2003), Cai and Xu (2008), and Cai and Xiao (2012) as special cases. In contrast to Koenker (2004), Galvao (2011), and Canay (2011) by requiring that both N and T go to infinity in their asymptotics, our model requires only N going to infinity with T possibly fixed. Actually, $T \geq 2$ is required, and indeed, it is an important assumption for identification; see Abrevaya and Dahl (2008) and Assumption A7 later. Also, different from Abrevaya and Dahl (2008) and Gamper-Rabindran et al. (2010), we use a partially varying-coefficient structure in the conditional quantile model to provide more flexibility in model specification than a linear model. Based on this empirical study, there are some novel findings. We find empirical evidence to support the absorptive capacity hypothesis, and furthermore, the host countries with fast economic growth can benefit more from FDI than ones with slow economic growth.

The rest of the paper is organized as follows. In Section 2, we introduce a partially varying-coefficient quantile panel data model with correlated random effects and propose a three-stage estimation procedure. Also, the asymptotic properties of our estimators are established. Furthermore, we propose a simple and easily implemented approach for testing the goodness-of-fit of a parametric model against model (1) and for constructing confidence intervals for parameters. In Section 3, a simulation study is conducted to examine the finite sample performance. Section 4 is devoted to reporting the empirical results of the cross-country panel data growth model. Section 5 concludes the paper and finally, all necessary notations and theoretical proofs are relegated to the appendices.

2. Econometric modeling

2.1. Model setup

In this paper, we consider the following partially varying-coefficient panel data quantile model with correlated random effects, in which there are both constant coefficients and varying coefficients. Let Y_{it}, a scalar dependent variable, be the observation on ith individual at time t for $1 \leq i \leq N$ and $1 \leq t \leq T$. The conditional quantile model is given by

$$Q_\tau(Y_{it} \mid U_{it}, \boldsymbol{X}_{it}, \alpha_i) = \boldsymbol{X}'_{it,1}\boldsymbol{\gamma}_\tau + \boldsymbol{X}'_{it,2}\boldsymbol{\beta}_\tau(U_{it}) + \alpha_i, \quad (1)$$

where $Q_\tau(Y_{it} \mid U_{it}, \boldsymbol{X}_{it}, \alpha_i)$ is the τth quantile of Y_{it} given U_{it}, $\boldsymbol{X}_{it}$, and α_i. Here, $\boldsymbol{X}_{it} = (\boldsymbol{X}'_{it,1}, \boldsymbol{X}'_{it,2})'$, where $\boldsymbol{X}_{it,1}$ and $\boldsymbol{X}_{it,2}$ are regressors with $K_1 \times 1$ and $K_2 \times 1$ dimensions, respectively, $\boldsymbol{\gamma}_\tau$ denotes a $K_1 \times 1$ vector of constant coefficients, $\boldsymbol{\beta}_\tau(U_{it})$ denotes a $K_2 \times 1$ vector of functional coefficients, U_{it} is an observable scalar smoothing variable,[5] and α_i is an individual effect. Model (1) allows for the dependence of the coefficients, both the constant coefficients and the functional coefficients, upon the quantile, but restricts α_i to a pure location shift effect, which is a common restriction in the quantile panel data literature.[6] If α_i is treated as a fixed effect, to estimate parameters and functionals in model (1), one could follow the ideas in Koenker (2004) by requiring both N and T go to infinity.[7] However, in this paper, T for our case is fixed so that we follow Abrevaya and Dahl (2008) and Gamper-Rabindran et al. (2010) and view the individual effect as a correlated random effect which is allowed to be correlated with covariates $\boldsymbol{X}_i = (\boldsymbol{X}'_{i1}, \ldots, \boldsymbol{X}'_{iT})'$ and $\{U_{it}\}_{t=1}^T$; that is,

$$\alpha_i = \alpha(\boldsymbol{X}_i, U_{i1}, \ldots, U_{iT}) + v_i, \quad (2)$$

where $\alpha(\cdot)$ is an unknown function of $\boldsymbol{X}_i$ and $\{U_{it}\}_{t=1}^T$, and v_i is a random error.

[5] For simplicity, we only consider the univariate case for the smoothing variable. The estimation procedure and asymptotic results still hold for the multivariate case with much complicated notation.

[6] Of course, it is possible and interesting to relax this restriction to allow α_i to depend on τ so that α_i in (1) becomes $\alpha_{i,\tau}$. But we conjecture that it is not an easy task. It warrants further investigation in the future.

[7] This is still an open research problem and it warrants for a further investigation in the future.

A fully nonparametric model of $\alpha(\cdot)$ may lead to the problem of the so called curse of dimensionality and become infeasible in practice. Compared with a linear projection in Chamberlain (1982) and Abrevaya and Dahl (2008), an additive model with functional coefficients can accommodate more flexibility. Thus, we approximate the unknown function $\alpha(\boldsymbol{X}_i, U_{i1}, \ldots, U_{iT})$ by a functional-coefficient model[8] such that

$$\alpha(\boldsymbol{X}_i, U_{i1}, \ldots, U_{iT}) = \sum_{t=1}^{T} \boldsymbol{X}_{it}' \boldsymbol{\delta}_t(U_{it}), \tag{3}$$

where $\boldsymbol{\delta}_t(U_{it})$ is a $K \times 1$ vector of unknown functional coefficients with $K = K_1 + K_2$.

Finally, in the case of estimating FDI effect on the economic growth in our empirical studies, the smoothing variable varies only across different individual units but keeps constant over time periods.[9] Therefore, in this paper, we focus on the simple case that $U_{it} = U_i$ for all $1 \le t \le T$. Model (1) can be rewritten as

$$Q_\tau(Y_{it} \mid U_i, \boldsymbol{X}_i, v_i) = \boldsymbol{X}_{it,1}' \boldsymbol{\gamma}_\tau + \boldsymbol{X}_{it,2}' \boldsymbol{\beta}_\tau(U_i) + \sum_{t=1}^{T} \boldsymbol{X}_{it}' \boldsymbol{\delta}_t(U_i) + v_i. \tag{4}$$

It is interesting to note that model (4) covers the following model for quantile regressions with measurement errors in dependent variable (EIV) as a special case. For simplicity, we consider the following simple quantile repression model for $T = 1$,

$$Q_\tau(Y_i | \boldsymbol{X}_i, v_i) = Q_\tau(\boldsymbol{X}_i) + v_i, \tag{5}$$

where $Q_\tau(\boldsymbol{X}_i)$ is the τth conditional quantile of $Y_i^0 = Y_i - v_i$ given $\boldsymbol{X}_i$. Here, Y_i^0 denotes true value but unobservable and Y_i is the observed value of Y_i^0 with the measurement error v_i. Model (5) might have many potential applications. For example, an empirical example of applying model (5) in labor economics is to study the heterogeneity of returns to education across conditional quantiles of the wage distribution; see Angrist et al. (2006) and Hausman et al. (2014) for details and the references therein.

2.2. Estimation procedures

2.2.1. Pooling regression strategy

From model (4), one can observe that the conditional quantile effects of $\boldsymbol{X}_{it}$ on Y_{it} are through two channels: a direct effect $\boldsymbol{\gamma}_\tau$ for constant coefficients and $\boldsymbol{\beta}_\tau(U_i)$ for varying coefficients, and an indirect effect $\boldsymbol{\delta}_t(U_i)$ working through the correlated random effects. Assuming that $T \ge 2$, to identify the direct effects $\boldsymbol{\gamma}_\tau$ and $\boldsymbol{\beta}_\tau(U_i)$, one has to estimate at least two conditional quantile models $Q_\tau(\boldsymbol{Y}_{it} \mid U_i, \boldsymbol{X}_i, v_i)$ and $Q_\tau(\boldsymbol{Y}_{is} \mid U_i, \boldsymbol{X}_i, v_i)$ given by

$$Q_\tau(Y_{it} | U_i, \boldsymbol{X}_i, v_i) = \boldsymbol{X}_{it,1}' [\boldsymbol{\gamma}_\tau + \boldsymbol{\delta}_{1t}(U_i)] + \boldsymbol{X}_{it,2}' [\boldsymbol{\beta}_\tau(U_i) + \boldsymbol{\delta}_{2t}(U_i)] + \sum_{l \ne t} \boldsymbol{X}_{il}' \boldsymbol{\delta}_l(U_i) + v_i$$

and

$$Q_\tau(Y_{is} | U_i, \boldsymbol{X}_i, v_i) = \boldsymbol{X}_{it,1}' \boldsymbol{\delta}_{1t}(U_i) + \boldsymbol{X}_{it,2}' \boldsymbol{\delta}_{2t}(U_i) + \boldsymbol{X}_{is,1}' \boldsymbol{\gamma}_\tau + \boldsymbol{X}_{is,2}' \boldsymbol{\beta}_\tau(U_i) + \sum_{l \ne t} \boldsymbol{X}_{il}' \boldsymbol{\delta}_l(U_i) + v_i,$$

respectively, where $t \ne s$, $\boldsymbol{\delta}_{1t}(U_i)$ is the vector containing the first K_1 components of $\boldsymbol{\delta}_t(U_i)$, and $\boldsymbol{\delta}_{2t}(U_i)$ is the vector for the last K_2 components of $\boldsymbol{\delta}_t(U_i)$. Hence, the estimates of $\boldsymbol{\gamma}_\tau$ and $\boldsymbol{\beta}(U_i)$ are respectively given by

$$\boldsymbol{\gamma}_\tau = \frac{\partial Q_\tau(\boldsymbol{Y}_{it} \mid U_i, \boldsymbol{X}_i, v_i)}{\partial \boldsymbol{X}_{it,1}} - \frac{\partial Q_\tau(\boldsymbol{Y}_{is} \mid U_i, \boldsymbol{X}_i, v_i)}{\partial \boldsymbol{X}_{it,1}},$$

and

$$\boldsymbol{\beta}_\tau(U_i) = \frac{\partial Q_\tau(\boldsymbol{Y}_{it} \mid U_i, \boldsymbol{X}_i, v_i)}{\partial \boldsymbol{X}_{it,2}} - \frac{\partial Q_\tau(\boldsymbol{Y}_{is} \mid U_i, \boldsymbol{X}_i, v_i)}{\partial \boldsymbol{X}_{it,2}}.$$

However, in order to avoid running two separating conditional quantile models, we adopt the pooling regression strategy as in Abrevaya and Dahl (2008) by stacking covariates. In view of model (4), $Q_\tau(\boldsymbol{Y}_{it} \mid U_i, \boldsymbol{X}_i, v_i)$ and $Q_\tau(\boldsymbol{Y}_{is} \mid U_i, \boldsymbol{X}_i, v_i)$ can be expressed as

$$Q_\tau(Y_{it} | U_i, \boldsymbol{X}_i, v_i) = \boldsymbol{X}_{it,1}' \boldsymbol{\gamma}_\tau + \boldsymbol{X}_{it,2}' \boldsymbol{\beta}_\tau(U_i) + \boldsymbol{X}_{i1}' \boldsymbol{\delta}_1(U_i) + \cdots + \boldsymbol{X}_{iT}' \boldsymbol{\delta}_T(U_i) + v_i,$$

8 As elaborated by Cai et al. (2006) and Cai (2010), a functional-coefficient model can be a good approximation to a fully nonparametric model, $g(X, Z) = \sum_{j=0}^{d} g_j(Z) X_j = X' g(Z)$.

9 When $U_{it_1} \ne U_{it_2} \ne U_i$ for any $t_1 \ne t_2$, two estimation approaches can be employed. We can apply the series estimation or adopt a single index method $U_i = \omega_1 U_{i1} + \cdots + \omega_T U_{iT}$ using the iterative backfitting method proposed by Fan et al. (2003).

and

$$Q_\tau(Y_{is}|U_i, \boldsymbol{X}_i, v_i) = \boldsymbol{X}'_{is,1}\boldsymbol{\gamma}_\tau + \boldsymbol{X}'_{is,2}\boldsymbol{\beta}_\tau(U_i) + \boldsymbol{X}'_{i1}\boldsymbol{\delta}_1(U_i) + \cdots + \boldsymbol{X}'_{iT}\boldsymbol{\delta}_T(U_i) + v_i.$$

Hence, we treat

$$\begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1T} \\ \vdots \\ Y_{i1} \\ \vdots \\ Y_{iT} \\ \vdots \\ Y_{N1} \\ \vdots \\ Y_{NT} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \boldsymbol{X}'_{11,1} & \boldsymbol{X}'_{11,2} & \boldsymbol{X}'_{11} & \cdots & \boldsymbol{X}'_{1T} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{X}'_{1T,1} & \boldsymbol{X}'_{1T,2} & \boldsymbol{X}'_{11} & \cdots & \boldsymbol{X}'_{1T} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{X}'_{i1,1} & \boldsymbol{X}'_{i1,2} & \boldsymbol{X}'_{i1} & \cdots & \boldsymbol{X}'_{iT} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{X}'_{iT,1} & \boldsymbol{X}'_{iT,2} & \boldsymbol{X}'_{i1} & \cdots & \boldsymbol{X}'_{iT} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{X}'_{N1,1} & \boldsymbol{X}'_{N1,2} & \boldsymbol{X}'_{N1} & \cdots & \boldsymbol{X}'_{NT} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{X}'_{NT,1} & \boldsymbol{X}'_{NT,2} & \boldsymbol{X}'_{N1} & \cdots & \boldsymbol{X}'_{NT} \end{pmatrix}$$

as the dependent variable and the right-side explanatory variables, respectively. This pooled regression directly estimates $\boldsymbol{\gamma}_\tau$ and $\boldsymbol{\theta}_\tau(U_i)$ with $\boldsymbol{\theta}_\tau(U_i) = (\boldsymbol{\beta}'_\tau(U_i), \boldsymbol{\delta}'_1(U_i), \ldots, \boldsymbol{\delta}'_T(U_i))'$. We now consider the following transformed model from (4),

$$Q_\tau(U_i, \boldsymbol{Z}_{it}, v_i) = \boldsymbol{Z}'_{it,1}\boldsymbol{\gamma}_\tau + \boldsymbol{Z}'_{it,2}\boldsymbol{\theta}_\tau(U_i) + v_i, \tag{6}$$

where $\boldsymbol{Z}_{it,1}$ denotes the corresponding variables in the first column in the above design matrix, $\boldsymbol{Z}_{it,2}$ represents those entries in the remaining columns, and $\boldsymbol{Z}_{it} = (\boldsymbol{Z}'_{it,1}, \boldsymbol{Z}'_{it,2})'$.

2.2.2. Quasi-likelihood function

For a conditional quantile regression model, according to Koenker and Bassett (1978), the estimation of parameters can be obtained by minimizing the following objective (loss) function

$$\hat{\theta}_{\mathrm{KB}} = \mathrm{argmin}_\theta L_{\mathrm{KB}}(\theta), \quad \text{where} \quad L_{\mathrm{KB}}(\theta) = \sum_{t=1}^{T} \rho_\tau(y_t - q_\tau(w_t, \theta)),$$

$q_\tau(w_t, \theta)$ is the conditional quantile regression function of y_t given w_t with unknown parameter θ, satisfying $P(y_t \leq q_\tau(w_t, \theta)|w_t) = \tau$, $\rho_\tau(x) = x(\tau - I_{x<0})$ is the so-called check function, and I_A is the indicator function of any set A. Komunjer (2005) generalized the estimation method of Koenker and Bassett (1978) by proposing a class of quasi-maximum likelihood estimations (QMLEs), which is $\hat{\theta}_{\mathrm{QMLE}}$, obtained by solving

$$\hat{\theta}_{\mathrm{QMLE}} = \mathrm{argmax}_\theta L_{\mathrm{QMLE}}(\theta), \quad \text{where} \quad L_{\mathrm{QMLE}} = \sum_{t=1}^{T} \ln l_t(y_t, q_\tau(w_t, \theta))$$

and $l_t(\cdot)$ is the conditional quasi-likelihood at the time t. As pointed out by Komunjer (2005), if $l_t(y_t, q_\tau(w_t, \theta))$ is taken to be $C(y_t, w_t)\exp(-\rho_\tau(y_t - q_\tau(w_t, \theta)))$ for some $C(y_t, w_t)$, the QMLE becomes the conventional estimator of Koenker and Bassett (1978).

We consider a class of integrated QMLEs for conditional quantile for the model defined in (6), obtained by solving the maximization of a quasi-likelihood function for the τth conditional quantile

$$\hat{\boldsymbol{\vartheta}} = \mathrm{argmax}_{\boldsymbol{\vartheta}} L_{\mathrm{I},\tau}(\boldsymbol{\vartheta}), \text{ where } L_{\mathrm{I},\tau}(\boldsymbol{\vartheta}) = \max_{\boldsymbol{\vartheta}} \sum_{i=1}^{N}\sum_{t=1}^{T} \ln l_{\mathrm{I},\tau}(Y_{it}, q_{\mathrm{I},\tau}(W_{it}, \boldsymbol{\vartheta}_i)), \tag{7}$$

where $l_{\mathrm{I},\tau}(Y_{it}, q_{\mathrm{I},\tau}(W_{it}, \boldsymbol{\vartheta}_i))$ is the integrated quasi-likelihood function for the τth conditional quantile on individual i at time t, $q_{\mathrm{I},\tau}(W_{it}, \boldsymbol{\vartheta}_i) = \boldsymbol{Z}'_{it,1}\boldsymbol{\gamma}_\tau + \boldsymbol{Z}'_{it,2}\boldsymbol{\theta}_\tau(U_i)$, $W_{it} = (U_i, \boldsymbol{Z}'_{it})'$, $\boldsymbol{\vartheta}_i = (\boldsymbol{\gamma}'_\tau, \boldsymbol{\theta}'_\tau(U_i))'$ and $\boldsymbol{\vartheta} = (\boldsymbol{\gamma}'_\tau, \boldsymbol{\theta}'_\tau(U_1), \ldots, \boldsymbol{\theta}'_\tau(U_N))'$. For simplicity, by assuming that v_i is iid as normal[10] with mean zero and variance σ^2, $l_{\mathrm{I},\tau}(y, q_{\mathrm{I}})$ is obtained from integrating the quasi-likelihood function for the τth conditional quantile with respect to v_i,

$$l_{\mathrm{I},\tau}(y, q_{\mathrm{I}}) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} l_\tau(y, q)e^{-\frac{v^2}{2\sigma^2}}\,dv,$$

[10] The normality assumption on v_i here is just for simplicity to obtain a close form for the quasi-likelihood (see (10) later). Of course, it can be relaxed but the quasi-likelihood function would be very complex. It would be very interesting to explore this issue in the future research.

where $q_I = q - v$, and $l_\tau(y, q)$ is the quasi-likelihood function for the τth conditional quantile. It is emphasized by Komunjer (2005) that different choices of $l_\tau(\cdot,\cdot)$ affect the asymptotic theory of the QMLE for quantile, similar to the case that different choices of likelihood function would affect the asymptotic theory of QMLE for mean model when the object of interest is the conditional mean. In this paper, for simplicity, we define $l_\tau(y, q)$ as

$$l_\tau(y, q) = e^{-\rho_\tau(y-q)}, \tag{8}$$

This definition makes $l_\tau(\cdot,\cdot)$ belonging to the so-called tick-exponential family defined by Komunjer (2005). Let $a = y - q_I = y - q + v$, then $y - q = a - v$. Substituting (8) into the integrated quasi-likelihood function for the τth conditional quantile, we have

$$l_\tau(a, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp[-\rho_\tau(a - v) - \frac{v^2}{2\sigma^2}]dv.$$

By a simple calculation, which can be found in Appendix F, we get

$$l_\tau(a, \sigma) = e^{-\rho_\tau(a)}\lambda_\tau(a, \sigma)(I_{a\geq 0} + e^{-a}I_{a<0}), \tag{9}$$

where $\lambda_\tau(a, \sigma) = e^{\frac{\tau^2\sigma^2}{2}}\Phi(\frac{a}{\sigma} - \tau\sigma) + e^{\frac{(\tau-1)^2\sigma^2}{2}}\Phi(-\frac{a}{\sigma} + (\tau - 1)\sigma)e^a$ and $\Phi(\cdot)$ is the standard normal distribution function. Thus,

$$\ln l_\tau(a, \sigma) = -\rho_\tau(a) + \ln(\lambda_\tau(a, \sigma)) + \ln(I_{a\geq 0} + e^{-a}I_{a<0}). \tag{10}$$

Clearly, the last two terms in (10) can be regarded as a penalty due to the randomness of v_i. Also, it is easy to show that the penalty approaches to zero as σ goes to zero. For Eq. (10), when $\sigma = 0$, it is exactly same as the case without v_i.

Therefore, the quasi-likelihood function $\mathrm{QL}_\tau(\boldsymbol{\vartheta}, \sigma)$ is given by

$$\begin{aligned}\mathrm{QL}_\tau(\boldsymbol{\vartheta}, \sigma) &= \sum_{i=1}^{N}\sum_{t=1}^{T} \ln l_\tau(a_{it}, \sigma) \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T}[-\rho_\tau(a_{it}) + \ln(\lambda_\tau(a_{it}, \sigma)) + \ln(I_{a_{it}\geq 0} + e^{-a_{it}}I_{a_{it}<0})].\end{aligned} \tag{11}$$

Hence, for a given $\tilde{\sigma}$ satisfying the following equation

$$\frac{\partial \mathrm{QL}_\tau(\boldsymbol{\vartheta}, \sigma)}{\partial\sigma}|_{\sigma=\tilde{\sigma}} = NT\tau^2\tilde{\sigma} + \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{-\phi\left(a_{\tau,it}(\tilde{\sigma})\right) + (1-2\tau)\tilde{\sigma}e^{\tilde{\sigma}a_{\tau-\frac{1}{2},it}(\tilde{\sigma})}\Phi\left(-a_{\tau-1,it}(\tilde{\sigma})\right)}{\Phi\left(a_{\tau,it}(\tilde{\sigma})\right) + e^{\tilde{\sigma}a_{\tau-\frac{1}{2},it}(\tilde{\sigma})}\Phi\left(-a_{\tau-1,it}(\tilde{\sigma})\right)} = 0 \tag{12}$$

with $a_{\tau,it}(\sigma) = a_{it}/\sigma - \tau\sigma$,[11] the QMLE of $\boldsymbol{\vartheta}$ is obtained by

$$\hat{\boldsymbol{\vartheta}} = (\hat{\boldsymbol{\gamma}}_\tau', \hat{\boldsymbol{\theta}}_\tau'(U_1), \ldots, \hat{\boldsymbol{\theta}}_\tau'(U_N))' = \arg\max_{\boldsymbol{\vartheta}} \mathrm{QL}_\tau(\boldsymbol{\vartheta}, \tilde{\sigma}). \tag{13}$$

Remark 1. We can simply estimate $\boldsymbol{\vartheta}$ and $\tilde{\sigma}$ by iterating (13) and (12) until convergence. Given an initial value of $\tilde{\sigma}$, we can estimate $\boldsymbol{\vartheta}$ by maximizing $\mathrm{QL}_\tau(\boldsymbol{\vartheta}, \tilde{\sigma})$. Then, we let $\tilde{\sigma} = \hat{\sigma}$ and iterate (13) and (12) until convergence.

2.2.3. Three-stage estimation procedure

To estimate the semiparametric model (6), we propose a three-stage estimation procedure to the proposed panel data model. At the first stage, we treat all coefficients as functional coefficients depending on U_i, such as $\boldsymbol{\gamma}_\tau = \boldsymbol{\gamma}_\tau(U_i)$. It is assumed throughout the paper that $\boldsymbol{\gamma}_\tau(\cdot)$ and $\boldsymbol{\theta}_\tau(\cdot)$ are both twice continuously differentiable, and then we apply the local constant approximations to $\boldsymbol{\gamma}_\tau(\cdot)$ and the local linear approximations to $\boldsymbol{\theta}_\tau(\cdot)$, respectively. Hence, model (6) is estimated as a fully functional-coefficient model and following Cai and Xu (2008), the localized quasi-likelihood function is given by

$$\max_{\boldsymbol{\gamma}_0,\boldsymbol{\theta}_0,\boldsymbol{\theta}_1}\sum_{i=1}^{N}\sum_{t=1}^{T}[-\rho_\tau(a_{it,1}) + \ln(\lambda_\tau(a_{it,1}, \tilde{\sigma})) + \ln(I_{a_{it,1}\geq 0} + e^{-a_{it,1}}I_{a_{it,1}<0})]K_h(U_i - u_0), \tag{14}$$

where $a_{it,1} = Y_{it} - \boldsymbol{Z}_{it,1}'\boldsymbol{\gamma}_0 - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_0 - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_1(U_i - u_0)$, $\boldsymbol{\gamma}_0 = \boldsymbol{\gamma}_\tau(u_0)$, $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_\tau(u_0)$, $\boldsymbol{\theta}_1 = \dot{\boldsymbol{\theta}}_\tau(u_0)$, $K_h(u) = K(u/h)/h$, and $K(\cdot)$ is the kernel function. Note that $\dot{A}$ and $\ddot{A}$ denote the first order and second order partial derivatives of A throughout the paper.

[11] The details of derivation of Eq. (12) can be found in Appendix F.

Since $\boldsymbol{\gamma}_\tau$ is a global parameter, in order to utilize all sample information to estimate $\boldsymbol{\gamma}_\tau$, at the second stage, we employ the average method to achieve the $\sqrt{N}$ consistent estimator of $\boldsymbol{\gamma}_\tau$, which is given by

$$\hat{\boldsymbol{\gamma}}_\tau = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\gamma}}_\tau(U_i). \tag{15}$$

Theorem 1 (see later) shows that indeed, $\hat{\boldsymbol{\gamma}}_\tau$ is a $\sqrt{N}$ consistent estimator.

Remark 2. First, it is worth pointing out that the well known profile least squares type of estimation approach (Robinson (1988) and Speckman (1988)) for classical semiparametric regression models may not be suitable to quantile setting due to lack of explicit normal equations. Secondly, the estimator $\hat{\boldsymbol{\gamma}}_\tau$ given in (15) has the advantage that it is easy to construct and also achieves the $\sqrt{N}$-rate of convergence (see Theorem 1 later). In addition to this simple estimator, other $\sqrt{N}$ consistent estimators of $\boldsymbol{\gamma}_\tau$ can be constructed. For example, to estimate the parameter $\boldsymbol{\gamma}_\tau$ without being overly influenced by the tail behavior of the distribution of U_i, one might use a trimming function $w_i = I_{\{U_i\in\mathcal{D}\}}$ with a compact subset $\mathcal{D}$ of $\mathbb{R}$; see Cai and Masry (2000) for details. Then, (15) becomes the weighted average estimator as

$$\tilde{\boldsymbol{\gamma}}_{w,\tau} = \left[\frac{1}{N}\sum_{i=1}^{N} w_i\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} w_i\hat{\boldsymbol{\gamma}}_\tau(U_i)\right].$$

Indeed, this type of estimator was considered by Lee (2003) for a partially linear quantile regression model. To estimate $\boldsymbol{\gamma}_\tau$ more efficiently, a general weighted average approach can be constructed as follows

$$\check{\boldsymbol{\gamma}}_{w,\tau} = \left[\frac{1}{N}\sum_{i=1}^{N} W(U_i)\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} W(U_i)\hat{\boldsymbol{\gamma}}_\tau(U_i)\right],$$

where $W(\cdot)$ is a weighting function (a symmetric matrix) which can be chosen optimally by minimizing the asymptotic variance of the estimator $\check{\boldsymbol{\gamma}}_{w,\tau}$; see Cai and Xiao (2012) for details. For simplicity, our focus is on $\hat{\boldsymbol{\gamma}}_\tau$ given in (15).

At the last stage, to estimate the varying coefficients, for a given $\sqrt{N}$-consistent estimator $\hat{\boldsymbol{\gamma}}_\tau$ of $\boldsymbol{\gamma}_\tau$ obtained from (15), we plug $\hat{\boldsymbol{\gamma}}_\tau$ into model (6) and obtain the partial residual denoted by $Y_{it}^* = Y_{it} - \boldsymbol{Z}_{it,1}'\hat{\boldsymbol{\gamma}}_\tau$. Hence, the functional coefficients can be estimated by using the local linear quantile estimation which is given by

$$\max_{\boldsymbol{\theta}_0,\boldsymbol{\theta}_1}\sum_{i=1}^{N}\sum_{t=1}^{T}[-\rho_\tau(a_{it,2}) + \ln(\lambda_\tau(a_{it,2},\tilde{\sigma})) + \ln(I_{a_{it,2}\geq 0} + e^{-a_{it,2}}I_{a_{it,2}<0})]K_h(U_i-u_0), \tag{16}$$

where $a_{it,2} = Y_{it}^* - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_0 - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_1(U_i-u_0)$. By moving u_0 along the domain of U_i, the entire estimated curve of the functional coefficient is obtained.

2.3. Asymptotic properties

This section provides asymptotic results of $\hat{\boldsymbol{\gamma}}_\tau$ and $\hat{\boldsymbol{\beta}}_\tau(u_0)$ defined in Section 2.2.3. To simplify the presentation here, all necessary notations and proofs are relegated to Appendix B. The following assumptions are used to establish the consistency and asymptotic normality of our estimators.

Assumptions:

A1. The series $\{U_i\}$ is iid. The series $\{\boldsymbol{Z}_{it}\}$ is iid across individual i, but can be correlated around t for fixed i. The series $\{v_i\}$ is iid $N(0,\sigma^2)$ and independent of $\{U_i, \boldsymbol{Z}_{it}\}$.

A2. The distribution of Y given U and $\boldsymbol{Z}$ has an everywhere positive Lebesgue density $f_{Y|U,\boldsymbol{Z}}(\cdot)$, which is bounded and satisfies the Lipschitz continuity condition.

A3. The kernel function $K(\cdot)$ is a symmetric bounded density with a bounded support region.

A4. The functional coefficients $\boldsymbol{\theta}(u_0)$ are two times continuously differentiable in a small neighborhood of u_0. The marginal density of U, $f_U(\cdot)$, is continuous with $f_U(u_0) > 0$. Assume all the variance–covariance matrices are positive-definite and continuously differentiable in a neighborhood of u_0.

A5. $E(\|\boldsymbol{Z}\|^{\delta^*}) < \infty$ with $\delta^* > 4$.

A6. The bandwidths h_1 at the first stage and h_2 at the third stage satisfy that $h_1 \to 0$, $h_2 \to 0$, $Nh_1 \to \infty$ and $Nh_2 \to \infty$ as $N \to \infty$.

A7. $T \geq 2$.

Assumption A1 assumes the data to be iid across individual i, but allows for arbitrary correlation around t for given i. The normality assumption of v_i can be relaxed by using some approximation approaches such as Laplace or saddle point approximation, or E-M algorithm, but the quasi-likelihood will be complicated since there is no close form of the quasi-likelihood like (9). Note that we do not exclude the heteroscedasticity dependence between individual effects and covariates through correlated random effects. Assumptions from A2 to A5 are standard in the nonparametric literature which impose some smooth and moment conditions on functionals involved. Assumptions A6 and A7 require that N go to infinity but T can be short. For the model with large T, some appropriate mixing condition should be imposed to restrict the dependence structure across t. Assumption A7 excludes the pure cross sectional data.

As mentioned above, a $\sqrt{N}$ consistent estimator of $\hat{\boldsymbol{\gamma}}_\tau$ at the second stage is constructed by using the average method defined in (15). The following theorem states its asymptotic normality result which can be obtained by using the U-statistic technique as in Powell et al. (1989).

Theorem 1. *Suppose that Assumptions A1–A7 hold, we have*

$$\sqrt{N}[\hat{\boldsymbol{\gamma}}_\tau - \boldsymbol{\gamma}_\tau - B_{\gamma,\tau}(\sigma)] \xrightarrow{d} N(0, \Sigma_{\gamma,\tau}(\sigma)),$$

where $B_{\gamma,\tau}(\sigma)$ and $\Sigma_{\gamma,\tau}(\sigma)$ are defined in Eqs. (23) and (24)*, respectively, in Appendix B, and $\xrightarrow{d}$ denotes the convergence in distribution. Furthermore, if $\sqrt{N}h_1^2 \to 0$, then, $\sqrt{N}[\hat{\boldsymbol{\gamma}}_\tau - \boldsymbol{\gamma}_\tau] \xrightarrow{d} N(0, \Sigma_{\gamma,\tau}(\sigma))$.*

Theorem 1 shows that the estimator $\hat{\boldsymbol{\gamma}}_\tau$ is $\sqrt{N}$ consistent and is asymptotically unbiased if the bandwidth at the first stage h_1 satisfies $\sqrt{N}h_1^2 \to 0$, which implies that the under-smooth at the first stage is needed. Therefore, it is common to assume that the first stage estimation is under-smoothed so that the asymptotic bias term disappears. Furthermore, we can see that $\Sigma_{\gamma,\tau}(\sigma)$ in the above theorem depends on both σ and T. Clearly, when T becomes larger, $\Sigma_{\gamma,\tau}(\sigma)$ becomes smaller. Also, when $\sigma = 0$, it reduces to $\Sigma_{\gamma,\tau} = \Sigma_{\gamma,\tau}(0)$. Thus, when $\sigma = 0$, the asymptotic normality of $\hat{\boldsymbol{\gamma}}_\tau$ reduces to the case without v_i.

At the last stage, the partial residuals Y_{it}^* are used to estimate $\hat{\boldsymbol{\theta}}_\tau(u_0)$. The following theorem depicts the asymptotic normality result of $\hat{\boldsymbol{\beta}}_\tau(u_0)$, where $\hat{\boldsymbol{\beta}}_\tau(u_0) = e_2'\hat{\boldsymbol{\theta}}_{0,\tau}(u_0)$ and e_2 is a selection matrix with ones in the first K_2 entries.

Theorem 2. *Suppose that Assumptions A1–A7 hold, given the $\sqrt{N}$ consistent estimator of $\boldsymbol{\gamma}_\tau$, we have*

$$\sqrt{Nh_2}[\hat{\boldsymbol{\beta}}_\tau(u_0) - \boldsymbol{\beta}_\tau(u_0) - B_{\beta,\tau}(u_0)] \to N(0, \Sigma_{\beta,\tau}(u_0, \sigma)),$$

where $B_{\beta,\tau}(u_0)$ and $\Sigma_{\beta,\tau}(u_0, \sigma)$ are defined in Eqs. (25) and (26)*, respectively.*

It is not surprising to see from Theorem 2 that the asymptotic bias term $B_{\beta,\tau}(u_0)$ does not depend on either σ or T but it comes only from the approximation bias in a nonparametric nature. Furthermore, we can see that similar to $\Sigma_{\gamma,\tau}(\sigma)$ above, $\Sigma_{\beta,\tau}(u_0, \sigma)$ does depend on both σ and T.

It is well documented that the selection of the optimal bandwidths is of importance in a nonparametric smoothing estimation. As far as we know, there is only few existing research focusing on the theoretical analysis of bandwidth selection in the field of nonparametric quantile regression. Recently, Li et al. (2015) made a great effort to this quite challenging work by using completely data driven cross-validation (CV) method. In the empirical analysis in Section 4, we adopt the CV method in Li et al. (2015) to choose the optimal bandwidth at the third stage. As for the bandwidth at the first stage, as aforementioned, we use the under-smoothed bandwidth. However, in the Monte Carlo simulation study, we follow the ad-hoc way to choose the bandwidths for time saving.

2.4. Inferences

Now, we turn to discussing how to test constancy on varying coefficients and accordingly construct confidence intervals. To make statistical inferences for $\boldsymbol{\gamma}_\tau$ and $\boldsymbol{\beta}_\tau(\cdot)$ in practice, we firstly need to obtain consistent covariance estimators of $\Sigma_{\gamma,\tau}(\sigma)$ and $\Sigma_{\beta,\tau}(u_0, \sigma)$, respectively. To this end, we need to estimate $\Omega_{z\dot{g}}(u_0, \sigma)$, $\Omega_{\tau,z}(u_0, \sigma)$, $\Omega_{\tau,z_{1t}}(u_0, \sigma)$, $\Omega_{z_2\dot{g}}(u_0, \sigma)$, $\Omega_{\tau,z_2}(u_0, \sigma)$ and $\Omega_{\tau,z_{1t,2}}(u_0, \sigma)$ consistently. Since the estimation of $\Omega_{z_2\dot{g}}(u_0, \sigma)$, $\Omega_{\tau,z_2}(u_0, \sigma)$ and $\Omega_{\tau,z_{1t,2}}(u_0, \sigma)$ is similar to $\Omega_{z\dot{g}}(u_0, \sigma)$, $\Omega_{\tau,z}(u_0, \sigma)$ and $\Omega_{\tau,z_{1t}}(u_0, \sigma)$, respectively, we here only focus on the latter to save notations.

We define

$$\hat{\Omega}_z(u_0) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' K_h(U_i - u_0),$$

$$\hat{\Omega}_{z_{1t}}(u_0) = (N(T-t))^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T-t} \boldsymbol{Z}_{is}\boldsymbol{Z}_{i(s+t)}' K_h(U_i - u_0),$$

$$\hat{\Omega}_{zg}(u_0, \sigma) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' \hat{m}_g(u_0, \boldsymbol{Z}_{it}, \sigma) K_h(U_i - u_0)$$

and

$$\hat{\Omega}_{z_{1t}g}(u_0, \sigma) = (N(T-t))^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T-t} \boldsymbol{Z}_{is}\boldsymbol{Z}_{i(s+t)}' \hat{m}_g(u_0, \boldsymbol{Z}_{i(s+t)}, \sigma) K_h(U_i - u_0),$$

where

$$\hat{m}_g(u,\boldsymbol{z},\sigma)=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}K_h(U_i-u,\boldsymbol{Z}_{it}-\boldsymbol{z})g(Y_{it}-\boldsymbol{z}_1'\boldsymbol{\gamma}_\tau-\boldsymbol{z}_2'\boldsymbol{\theta}_\tau(u),\sigma)}{\sum_{i=1}^{N}\sum_{t=1}^{T}K_h(U_i-u,\boldsymbol{Z}_{it}-\boldsymbol{z})}.$$

It can be easily shown that $\hat{\Omega}_z(u_0)=f_U(u_0)\Omega_z(u_0)+o_p(1)$, $\hat{\Omega}_{zg}(u_0,\sigma)=f_U(u_0)\Omega_{zg}(u_0,\sigma)+o_p(1)$, $\hat{\Omega}_{z_{1t}}(u_0)=f_U(u_0)\Omega_{z_{1t}}(u_0)+o_p(1)$ and $\hat{\Omega}_{z_{1t}g}(u_0,\sigma)=f_U(u_0)\Omega_{z_{1t}g}(u_0,\sigma)+o_p(1)$. Similarly, we can get the consistent estimators of $\Omega_{zg^2}(u_0,\sigma)$, $\Omega_{z_{1t}g_{1t}}(u_0,\sigma)$ and $\Omega_{z\dot{g}}(u_0,\sigma)$. Then, $\hat{\Omega}_{\tau,z}(u_0,\sigma)=\tau^2\hat{\Omega}_z(u_0)-2\tau\hat{\Omega}_{zg}(u_0,\sigma)+\hat{\Omega}_{zg^2}(u_0,\sigma)$ and $\hat{\Omega}_{\tau,z_{1t}}(u_0,\sigma)=\tau^2\hat{\Omega}_{z_{1t}}(u_0)-2\tau\hat{\Omega}_{z_{1t}g}(u_0,\sigma)+\hat{\Omega}_{z_{1t}g_{1t}}(u_0,\sigma)$. Next, the consistent covariance estimator of $\Sigma_{\gamma,\tau}(\sigma)$ can be given by

$$e_1'N^{-1}\sum_{i=1}^{N}\hat{\Omega}_{z\dot{g}}^{-1}(U_i,\sigma)[\frac{1}{T}\hat{\Omega}_{\tau,z}(U_i,\sigma)+\sum_{t=2}^{T}\frac{2(T-t+1)}{T^2}\hat{\Omega}_{\tau,z_{1t}}(U_i,\sigma)]\hat{\Omega}_{z\dot{g}}^{-1}(U_i,\sigma)e_1.$$

Finally, the consistent estimator of $\Sigma_{\beta,\tau}(u_0,\sigma)$ can be constructed accordingly in an obvious manner.

In empirical studies, it is of importance to test the constancy of the varying coefficients. Following Cai and Xiao (2012), a null hypothesis is given by

$$H_0:\beta_\tau(u_j)=\beta_\tau\quad\text{for some}\quad\{u_j\}_{j=1}^{q},$$

where $\{u_j\}_{j=1}^{q}$ denotes a set of distinct points within the domain of U_i. Cai and Xiao (2012) provided some comments on the choice of $\{u_j\}_{j=1}^{q}$ and q in practice. Under the null hypothesis, a simple and easily implemented test statistic can be constructed as follows

$$T_N=\sum_{1\le j\le q}\|\sqrt{Nh_2}\hat{\Sigma}_{\beta,\tau}(u_j,\sigma)^{-1/2}(\hat{\boldsymbol{\beta}}_\tau(u_j)-\hat{\boldsymbol{\beta}}_\tau)\|^2\to\chi^2_{qK_2} \tag{17}$$

where $\chi^2_{qK_2}$ is a chi-squared random variable with qK_2 degrees of freedom. Thus, the null is rejected if T_N is too large. Note that the proposed test statistic T_N in (17) is slightly different from that in Cai and Xiao (2012) by using the maximum instead of summation in (17). Of course, other types of test statistics may be constructed and this is left as future research topics.

3. A Monte Carlo simulation study

In this section, we conduct Monte Carlo simulations to demonstrate the finite sample performance of the proposed estimators for both constant and functional coefficients. We consider the following data generating process

$$Y_{it}=\gamma_0+X_{it,1}\gamma_1+X_{it,2}\beta(U_i)+\sum_{t=1}^{T}X_{it,2}\delta_t(U_i)+v_i+(0.5+0.3X_{it,1}+0.4X_{it,2})\varepsilon_{it} \tag{18}$$

with $T=2$, where U_i is generated from iid $U(-2.5,2.5)$, $X_{it,1}$ and $X_{it,2}$ are respectively generated from iid $U(0,3)$ and $U(0,2)$, and ε_{it} is generated from iid $N(0,1)$. The constant coefficients above are set by $\gamma_0=2$ and $\gamma_1=-1.5$, respectively. The functional coefficients are defined as $\beta(u)=0.5\cos(2u)+u/3$, $\delta_1(u)=\sin(1.5u)$ and $\delta_2(u)=1.5e^{-u^2}-0.75$. To check the robustness, v_i is generated from iid normal and iid non-normal, respectively. For the iid normal case, we choose $\mu=0$ and $\sigma=0.1,\ 0.3,\ 0.5$. For the iid non-normal case, we choose Laplace with $\mu=0$ and $\sigma=0.1$.

To measure the performance of $\hat{\gamma}_{j,\tau}$ for $0\le j\le 1$ and $\hat{\beta}_\tau(\cdot)$, we use the mean absolute deviation errors (MADE) of the estimators, which is defined by

$$\text{MADE}(\hat{\beta}_\tau(\cdot))=\frac{1}{n_0}\sum_{l=1}^{n_0}|\hat{\beta}_\tau(u_l)-\beta_\tau(u_l)|,$$

where $\{u_l\}_{l=1}^{n_0}$ are the grid points within the domain of U_i, and

$$\text{MADE}(\hat{\gamma}_{j,\tau})=|\hat{\gamma}_{j,\tau}-\gamma_{j,\tau}|$$

for $0\le j\le 1$.

We consider three different sample sizes, $N=200$, 500 and 1000, respectively. For each given sample size, we repeat simulations by 500 times to calculate the MADE values. We compare the estimation results using different bandwidths, for example, $h_1=5N^{-2/5}$ (under-smooth) and $h_2=cN^{-1/5}$, where c is chosen from 1.5, 1.7, 2, 2.2, 2.5, 2.7, 3.0, $\cdots$. From simulation results we find that the estimation of constant coefficients is not sensitive to the choice of the bandwidth when the first stage is under-smoothed, and the estimation of $\beta_\tau(\cdot)$ is quite stable when the bandwidth is chosen within a reasonable range. The optimal bandwidth for estimating functional coefficient $\beta_\tau(\cdot)$ in our experiments is about $h_2=2.7N^{-1/5}$.

The simulation results of the median and standard deviation (denoted by SD) in parentheses for both estimators are summarized in Table 1. From Table 1, we can observe that the medians of 500 MADE values in all settings decrease significantly as N increases. For the case of $v_i\sim$ *normal* with $\mu=0$ and $\sigma=0.3$, when the sample size increases from

Table 1
The median and SD of the MADE values for $\hat{\gamma}_{0,\tau}$, $\hat{\gamma}_{1,\tau}$ and $\hat{\beta}_{\tau}(\cdot)$.

	$\tau = 0.15$			$\tau = 0.5$			$\tau = 0.75$		
	$\gamma_{0,\tau}$	$\gamma_{1,\tau}$	$\beta_{\tau}(\cdot)$	$\gamma_{0,\tau}$	$\gamma_{1,\tau}$	$\beta_{\tau}(\cdot)$	$\gamma_{0,\tau}$	$\gamma_{1,\tau}$	$\beta_{\tau}(\cdot)$
Case 1: $v_i \sim$ *Normal* with $\mu = 0$ and $\sigma = 0.1$									
200	0.1872	0.1154	0.2883	0.1488	0.0950	0.2330	0.1839	0.1113	0.2495
	(0.1724)	(0.1142)	(0.0827)	(0.1451)	(0.0914)	(0.0689)	(0.1594)	(0.0995)	(0.0727)
500	0.1177	0.0715	0.1917	0.0957	0.0603	0.1603	0.0961	0.0629	0.1720
	(0.1055)	(0.0631)	(0.0499)	(0.0819)	(0.0497)	(0.0450)	(0.0911)	(0.0533)	(0.0514)
1000	0.0885	0.0508	0.1464	0.0675	0.0404	0.1221	0.0674	0.0444	0.1296
	(0.0738)	(0.0449)	(0.0385)	(0.0625)	(0.0366)	(0.0315)	(0.0678)	(0.0406)	(0.0324)
Case 2: $v_i \sim$ *Normal* with $\mu = 0$ and $\sigma = 0.3$									
200	0.2152	0.1089	0.2716	0.1609	0.0900	0.2244	0.1742	0.1006	0.2379
	(0.1931)	(0.1044)	(0.0828)	(0.1395)	(0.0828)	(0.0660)	(0.1608)	(0.0950)	(0.0722)
500	0.1512	0.0790	0.1925	0.0969	0.0560	0.1578	0.1203	0.0586	0.1632
	(0.1193)	(0.0662)	(0.0574)	(0.0821)	(0.0542)	(0.0456)	(0.1002)	(0.0615)	(0.0476)
1000	0.1293	0.0581	0.1376	0.0667	0.0363	0.1172	0.0905	0.0444	0.1268
	(0.0974)	(0.0455)	(0.0398)	(0.0565)	(0.0338)	(0.0329)	(0.0748)	(0.0399)	(0.0332)
Case 3: $v_i \sim$ *Normal* with $\mu = 0$ and $\sigma = 0.5$									
200	0.3005	0.1250	0.2701	0.1695	0.0960	0.2169	0.2189	0.1047	0.2326
	(0.2406)	(0.1185)	(0.0898)	(0.1463)	(0.0913)	(0.0669)	(0.1856)	(0.0980)	(0.0750)
500	0.2583	0.0871	0.1934	0.1020	0.0593	0.1571	0.1697	0.0653	0.1615
	(0.1624)	(0.0771)	(0.0537)	(0.0842)	(0.0534)	(0.0418)	(0.1302)	(0.0588)	(0.0456)
1000	0.2536	0.0719	0.1510	0.0746	0.0466	0.1172	0.1700	0.0603	0.1263
	(0.1185)	(0.0582)	(0.0424)	(0.0617)	(0.0382)	(0.0324)	(0.1028)	(0.0463)	(0.0331)
Case 4: $v_i \sim$ *Laplace* with $\mu = 0$ and $\sigma = 0.1$									
200	0.1913	0.1179	0.2791	0.1525	0.0958	0.2296	0.1705	0.1035	0.2531
	(0.1669)	(0.1059)	(0.0807)	(0.1262)	(0.0849)	(0.0721)	(0.1489)	(0.0959)	(0.0772)
500	0.1174	0.0726	0.1895	0.0954	0.0574	0.1656	0.1108	0.0637	0.1734
	(0.1055)	(0.0644)	(0.0537)	(0.0877)	(0.0526)	(0.0450)	(0.0980)	(0.0583)	(0.0497)
1000	0.0836	0.0511	0.1458	0.0688	0.0469	0.1220	0.0763	0.0451	0.1272
	(0.0718)	(0.0494)	(0.0385)	(0.0601)	(0.0397)	(0.0317)	(0.0676)	(0.0422)	(0.0342)

200 to 1000, the medians of MADE values for $\hat{\gamma}_{0,0.15}$, $\hat{\gamma}_{1,0.15}$ and $\hat{\beta}_{0.15}(\cdot)$ all shrink quickly, from 0.2152 to 0.1293, from 0.1089 to 0.0581, and from 0.2716 to 0.1376, respectively. The standard deviations also shrink quickly when the sample size is enlarged. For example, for $\hat{\gamma}_{0,0.15}$, the standard deviation shrinks from 0.1931 to 0.0974, and for $\hat{\gamma}_{1,0.15}$ and $\hat{\beta}_{0.15}(\cdot)$, they decrease from 0.1044 to 0.0455 and from 0.0828 to 0.0398, respectively. Similar results can also be observed at the median, $\tau = 0.5$, and at the upper quantile, $\tau = 0.75$. We observe similar results for other normal cases. For the non-normal experiment in Case 4, we also observe similar patterns as the sample size increases, although the mean absolute deviation errors and standard deviations are a little bit larger than those in Case 1. All results are in line with our asymptotic theory which implies that our proposed estimators are indeed consistent. Compared with the estimation of $\hat{\beta}_{\tau}(\cdot)$, the shrinkage speed of the estimation of $\hat{\gamma}_{1,\tau}$ is relatively fast, which is consistent with the theoretical results in the previous sections. For the normal cases, the medians of 500 MADE values for $\hat{\gamma}_{1,\tau}$ increase as the value of σ increases at almost every quantile level when the sample size is reasonably large, but the mixed results can be observed for $\hat{\beta}_{\tau}(\cdot)$, which is also consistent with the theoretical results that the bias of $\hat{\beta}_{\tau}(\cdot)$ does not depend on σ.

4. Modeling the effect of FDI on economic growth

4.1. Empirical models

A typical linear model in empirical studies to estimate the impact of FDI on economic growth, such as Kottaridi and Stengos (2010), is given by

$$y_{it} = \alpha_i + \beta_1(\mathrm{FDI}/Y)_{it} + \beta_2 \log(\mathrm{DI}/Y)_{it} + \beta_3 n_{it} + \beta_4 h_{it} + \varepsilon_{it}, \tag{19}$$

where y_{it} denotes the growth rate of GDP per capita in the country or region i during the period t, α_i is the individual effect used to control the unobserved country-specific heterogeneity, n_{it} is the logarithm of population growth rate, h_{it} is the human capital, and ε_{it} is the random error. Moreover, the FDI and DI in (19) refer to foreign direct investment and domestic investment respectively and Y represents the total output. Hence, $(\mathrm{FDI}/Y)_{it}$ denotes the average ratio between the FDI and the total output during the period t in country i and $(\mathrm{DI}/Y)_{it}$ is defined in the same fashion for the domestic investment. To allow the possible joint effect of FDI and human capital, some literatures considered to add an interacted term between FDI and human capital into the empirical growth model, then (19) becomes

$$y_{it} = \alpha_i + \beta_1(\mathrm{FDI}/Y)_{it} + \beta_2 \log(\mathrm{DI}/Y)_{it} + \beta_3 n_{it} + \beta_4 h_{it} + \beta_5((\mathrm{FDI}/Y)_{it} \times h_{it}) + \varepsilon_{it}, \tag{20}$$

see Kottaridi and Stengos (2010) and the references therein.

Since the majority of the literature realized that the effect of FDI on the economic growth depends on the absorptive capacity in host countries and the initial GDP per capita is one of the most important indicators to reflect the initial conditions and the absorptive capacity in the host country; see Hansen (2000), Nunnenkamp (2004), and among others, we hereby propose a partially varying-coefficient model which allows the effect of FDI on the economic growth to depend on the initial GDP per capita in the host country. Hence, our first econometric model is the following conditional semiparametric mean model given by

$$y_{it} = \alpha_i + \beta_1(U_i)(\mathrm{FDI}/Y)_{it} + \beta_2 \log(\mathrm{DI}/Y)_{it} + \beta_3 n_{it} + \beta_4 h_{it} + \beta_5((\mathrm{FDI}/Y)_{it} \times h_{it}) + \varepsilon_{it}, \tag{21}$$

where U_i is the logarithm of initial GDP per capita in country i and $\beta_1(U_i)$ is the varying coefficient over the logarithm of initial GDP per capita U_i. Therefore, model (21) has an ability to characterize how FDI may have different nonlinear effects on economic growth among host countries with different initial conditions. Note that model (21) is new in the growth literature on studying the effect of FDI on the economic growth even under the conditional mean framework.

Moreover, as we discussed in the introduction, the conditional mean model in (21) is usually insufficient to control the heterogeneity among countries although (21) has some nice properties. The existing literature dealt with the aforementioned issues by simply looking at sub-samples. Instead, in this paper, we propose adopting a quantile regression approach to investigate the impact of FDI on economic growth. Our method is capable of dealing with heterogeneity among countries by allowing different quantiles to have different empirical growth equations, and at the same time, we can avoid splitting the sample. Different from the mean model, another advance of considering the quantile model is to effectively characterize the heterogeneity effect of FDI in different groups of countries, for example, the economy fast growing countries (upper quantile) and the economy slowly growing countries (lower quantile).

Finally, we consider the following conditional quantile model,

$$\begin{aligned} \boldsymbol{Q}_\tau(y_{it} \mid U_i, \boldsymbol{X}_{it}, \alpha_i) = {} & \alpha_i + \beta_{1,\tau}(U_i)(\mathrm{FDI}/Y)_{it} + \beta_{2,\tau} \log(\mathrm{DI}/Y)_{it} \\ & + \beta_{3,\tau} n_{it} + \beta_{4,\tau} h_{it} + \beta_{5,\tau}((\mathrm{FDI}/Y)_{it} \times h_{it}), \end{aligned} \tag{22}$$

which can be regarded as a special case of model (1). Imposing the correlated random effect assumption in (3), we can derive the transformed conditional quantile regression model in (6). Therefore, the three-stage estimation procedures described in Section 2 can be applied here.

4.2. The data and empirical results

Our data set includes 95 countries or regions from 1970 to 1999. To smooth the yearly fluctuations in aggregate economic variables, we take five-year averages by following the convention in the empirical growth literature; see Maasoumi et al. (2007), Durlauf et al. (2008), and Kottaridi and Stengos (2010). The population growth is computed by the average annual growth rate in each period, the human capital is measured as mean years of schooling in each period, and the domestic investment refers to the average of the domestic gross fixed capital formation measured by the US dollars in 2000 constant values. We measure the initial GDP by the GDP per capita of each country/region at the beginning year of each decade in constant 2000 US dollars.[12] All the above data are available to be downloaded from World Development Indicators (WDI). The FDI flows, in constant 2000 US dollars, are taken from United Nations Conference on Trade and Development (UNCTAD). The full list of countries and regions can be found in Table 5 in Appendix A.

Firstly, we consider the classical linear regression model in (20). Table 2 presents corresponding estimation results, including coefficient estimates, standard deviations, t-values and p-values from Column 2 to Column 5, respectively. The estimate of β_1 is about 0.56, which is positive and significant with a p-value of 0.027. On average, the linear conditional mean model reports a mild positive effect of FDI on promoting the economic growth. Compared to the growth effect of FDI, Table 2 reports a larger effect of domestic investments on the economic growth, which is about 2.72 and highly significant with the p-value of 0.009. The effect of population growth (β_3) is also positive and significant, with an estimate around 0.65. However, other estimates (β_4 and β_5) are not significant.

Next, we move to the partially varying-coefficient conditional mean model in (21). Compared to the linear model in (20), we now allow $\beta_1(\cdot)$ to depend on the initial conditions. Fig. 1 and Table 3 present the corresponding estimation results. The solid line in Fig. 1 represents the nonparametric estimate of the varying coefficient $\beta_1(\cdot)$ along various values of initial GDP, and the shaded area is the corresponding 90% pointwise confidence intervals with the asymptotic bias ignored. The nonparametric estimate shows a mild but clear pattern that $\beta_1(\cdot)$ increases as the initial GDP improves, which is in line with the hypothesis of the absorptive capacity. The range of the estimated values of the varying coefficient is between 0.9 and 1.6 for different initial GDPs, much larger than 0.56, the estimated value of the linear model. Table 3 reports the estimates of constant coefficients in (21), which are quite different from the corresponding estimation results in Table 2. For example,

[12] We combine three decades, from 1970 to 1979 (69 countries), from 1980 to 1989 (93 countries), and from 1990 to 1999 (95 countries), and then obtain a panel of 514 observations with $N = 257$ and $T = 2$.

Table 2
Empirical results of a linear conditional mean model in (20).

Mean Model	Coefficient	Standard deviation	t-value	p-value
β_1	0.5589	0.2513	2.2248	0.0270 *
β_2	2.7180	1.0374	2.6206	0.0093 **
β_3	0.6483	0.2575	2.5183	0.0124 *
β_4	−0.0290	0.4269	−0.0682	0.9458
β_5	−0.0359	0.0375	−0.9561	0.3401

Table 3
Constant coefficients of a partial linear conditional mean model in (21).

Mean Model	Coefficient	Standard deviation	t-value	p-value
β_2	3.8100	0.1321	28.8326	0.0000 ***
β_3	−1.1838	0.3357	−3.5268	0.0004 ***
β_4	0.1728	0.0268	6.4519	0.0000 ***
β_5	−0.1753	0.0112	−15.5394	0.0000 ***

Table 4
p-values of testing constancy.

τ	$\beta_{1,\tau}$	$\beta_{2,\tau}$	$\beta_{3,\tau}$	$\beta_{4,\tau}$	$\beta_{5,\tau}$
0.15	0.0000	0.9999	0.9999	0.9999	0.9999
0.50	0.0000	0.9999	0.9999	0.9999	0.9999
0.75	0.0000	0.9999	0.9999	0.9999	0.9999

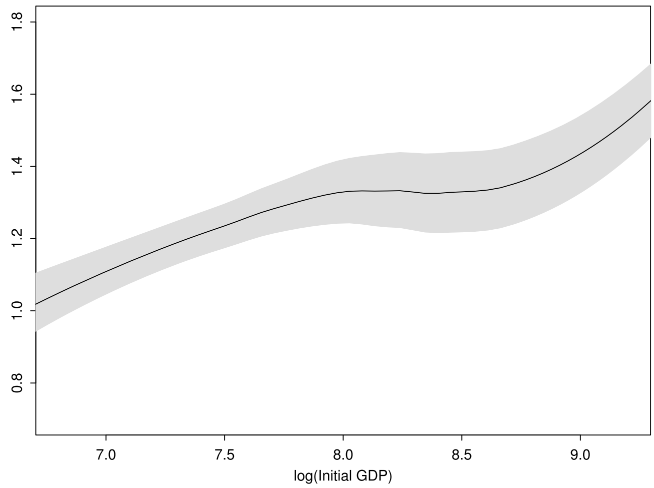

Fig. 1. Estimated curve of functional coefficient $\beta_1(\cdot)$ in model (21) together with the pointwise 90% confidence interval with the bias ignored.

in Table 3, the estimate of β_2 is now 3.81 instead of 2.72. The impact of population growth rate on the economic growth now becomes significantly negative with an estimate of -1.18. Moreover, both β_4 and β_5 become significant in Table 3. The estimate of the impact of human capital (not considering the interactive effect) is positive with a value of 0.17 and the estimate of the interacted term is -0.18. We attribute the different estimation results to the existence of nonlinearity in the regression model.

Finally, we consider the partially varying-coefficient quantile model in (22). We firstly conduct a constancy test as in Section 2.4 to testing whether or not all coefficients vary with the initial GDP at different quantile levels. The testing results are summarized in Table 4. It turns out that the constancy test only strongly rejects the null of constancy of $\beta_{1,\tau}(\cdot)$ but not other coefficients. All these results support our model setup in (22), where the coefficient of FDI depends on the initial conditions but others remain constant.

Fig. 2 presents the estimates of all four constant coefficients $\beta_{j,\tau}$ for $2 \le j \le 5$ under different quantile levels as $\tau = 0.15$, $0.25, 0.35, 0.45, \cdots, 0.75, 0.85$. The horizontal axis represents different quantiles and the vertical axis measures the estimated value of $\beta_{j,\tau}$. The curves in solid line denote the estimates under different quantiles and the areas in dark gray color are corresponding 90% confidence intervals. The horizontal solid lines denote the estimates under conditional mean model and the areas in light gray color are corresponding 90% confidence intervals. Except the estimated values of $\beta_{3,\tau}$ in the upper left panel in Fig. 2, most estimated values are outside the 90% confidence intervals of the conditional mean estimates, implying

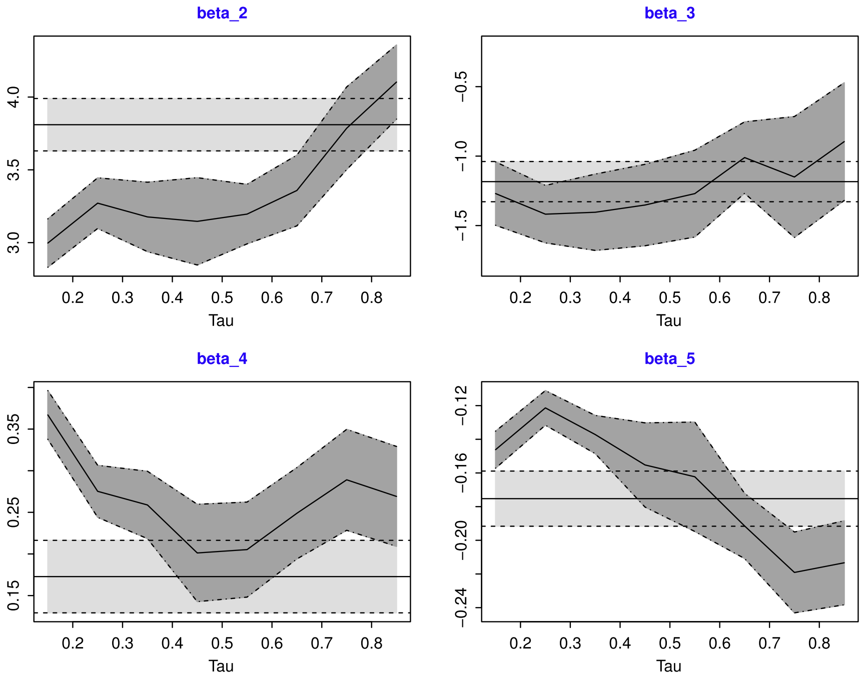

Fig. 2. Estimated values of constant coefficients $\beta_{j,\tau}$ in model (22) for $2 \leq j \leq 5$ and β_j in model (21) and their corresponding 90% confidence intervals.

that indeed, $\beta_{j,\tau}$ changes over τ and the conditional mean model might be inadequate to catch the heterogeneity effect. We observe that the estimated values of $\beta_{2,\tau}$ increase with τ when τ is bigger than 0.35 but the estimated values of $\beta_{5,\tau}$ decrease with τ when τ is in the range from 0.25 to 0.75. The estimated values of $\beta_{4,\tau}$ decrease with τ when $\tau < 0.5$ and increase with τ when $\tau > 0.5$. Moreover, the estimated values of $\beta_{2,\tau}$ and $\beta_{4,\tau}$ are all positive but the estimated values of $\beta_{3,\tau}$ and $\beta_{5,\tau}$ are all negative. Hence, generally speaking, we find a clear evidence that domestic investments and human capital (not considering the interactive effect) have positive effects on the economic growth, while the effect of domestic investments is larger and increases more significantly in countries or regions with better economic growth performance than those with poor growth performance. Moreover, the effect of human capital (not considering the interaction effect) shows a U-shape across countries or regions from poorer economic growth performance to better economic growth performance. However, the interaction effect between FDI and human capita is significantly negative for all quantiles, ranging from −0.22 to −0.14.

The nonparametric estimates of functional coefficient $\beta_{1,\tau}(\cdot)$ with the upper ($\tau = 0.75$, in the dashed line) and lower ($\tau = 0.15$, in the solid line) quantiles are demonstrated in Fig. 3. The horizontal axis measures different values of log of initial GDP, U_i, and the vertical axis measures the nonparametric estimated values of $\beta_{1,\tau}(\cdot)$. The shaded areas represent the 90% confidence intervals of $\hat{\beta}_{1,\tau}(\cdot)$ with the asymptotic bias ignored. We observe that the estimated values of $\beta_{1,\tau}(\cdot)$ at the upper quantile are significantly higher than those at the lower quantile uniformly over the values of initial GDPs. In general, our empirical findings support the hypothesis of absorptive capacity. The initial conditions really matter for host countries to benefit from adopting foreign direct investments. At the upper quantile, the estimated values of $\beta_{1,\tau}(\cdot)$ generally increase with the value of initial GDPs, and furthermore, the tendency of increase speeds up when $U_i > 8.2$. Note that $U_i = 8.2$ refers to a country or region with an initial GDP per capita about 3640.95 USD in constant 2000 dollars. However, at the lower quantile, although the estimated curve has an overall positive slope, it becomes almost flat when U_i is larger than 8.2 for host countries.

5. Conclusion

Quantile panel data models have gained a lot of attentions in the literature during recent years. In this paper, we propose a partially varying-coefficient quantile panel data model with correlated random effects. Compared to quantile panel data models with fixed effect, our estimation assumes only large N and short T, while the latter requires that both N and T go to infinity. In our semiparametric model, we allow some coefficients to vary with other economic variables while others keep constant. This novel semiparametric quantile panel data model is applied to estimating the impact of FDI on the economic growth.

There are several issues still worth of further studies. First, it is reasonable to allow for cross sectional dependence in the current model. In the literature of conditional mean models, some methods have been developed to deal with cross sectional

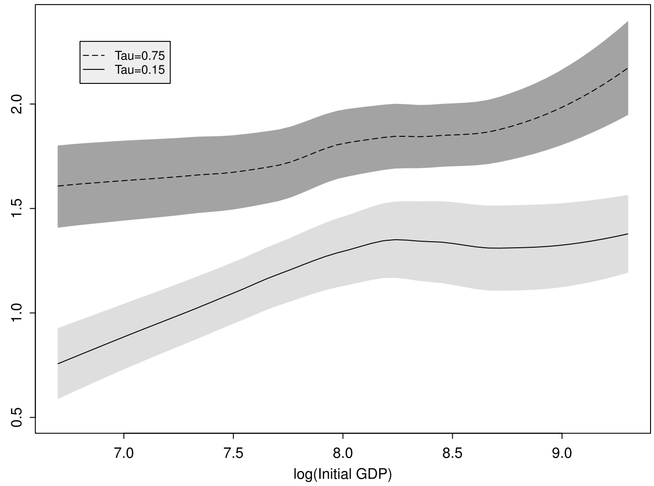

Fig. 3. Estimated curves of functional coefficient $\beta_{1,\tau}(\cdot)$ in model (22) for $\tau = 0.15$ (solid line) and $\tau = 0.75$ (dashed line) and their corresponding 90% pointwise confidence intervals with the bias ignored.

Table 5
Countries and regions in the empirical data set.

Algeria	Australia	Austria	Bahrain
Bangladesh	Barbados	Belgium	Benin
Bolivia	Botswana	Brazil	Cameroon
Canada	Central African Republic	Chile	China
Colombia	Congo, Rep.	Costa Rica	Cyprus
Denmark	Dominican Republic	Ecuador	Egypt, Arab Rep.
El Salvador	Fiji	Finland	France
Gambia	Germany	Ghana	Greece
Guatemala	Guyana	Honduras	Hong Kong SAR, China
Hungary	Iceland	India	Indonesia
Iran, Islamic Rep.	Ireland	Israel	Italy
Jamaica	Japan	Jordan	Kenya
Korea, Rep.	Lesotho	Malawi	Malaysia
Mali	Malta	Mauritius	Mexico
Mozambique	Nepal	Netherlands	New Zealand
Nicaragua	Niger	Norway	Pakistan
Panama	Papua New Guinea	Paraguay	Peru
Philippines	Poland	Portugal	Rwanda
Senegal	Sierra Leone	Singapore	South Africa
Spain	Sri Lanka	Sudan	Swaziland
Sweden	Switzerland	Syrian Arab Republic	Thailand
Togo	Trinidad and Tobago	Tunisia	Turkey
Uganda	United Kingdom	United States	Uruguay
Venezuela, RB	Zambia	Zimbabwe	

dependence, for example, using the factor structure or the interactive effect. However, due to the nature of conditional quantile model, it is not obvious to extend these under the quantile setup. Second, it is also interesting to address a dynamic structure and endogeneity issue in conditional quantile panel data models. We leave these important issues as future research topics.

Appendix A. Table of countries and regions

See Table 5.

Appendix B. Notations and definitions

All notations and definitions given here will be used in the following sections. We define $\mu_j = \int_{-\infty}^{\infty} u^j K(u)du$ and $\nu_j = \int_{-\infty}^{\infty} u^j K^2(u)du$ with $j \geq 0$, and denote h_1 and h_2 to be the bandwidths used at the first and third stages, respectively.

Let $e_1' = (\boldsymbol{I}_{K_1^*}, \boldsymbol{0}_{K_1^* \times K_2^*})$ and $e_2' = (\boldsymbol{I}_{K_2}, \boldsymbol{0}_{K_2 \times KT})$, where $K^* = K_1^* + K_2^*$, $K_1^* = K_1$ and $K_2^* = K_2 + KT$. Denote $f_U(\cdot)$ by the marginal density of U. Let $g_\tau(a, \sigma) = \partial \ln(\lambda_\tau(a, \sigma))/\partial a$[13], $b_{it,1}(U_i) = Y_{it} - \boldsymbol{Z}_{it,1}'\boldsymbol{\gamma}_\tau - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i)$ and $b_{it,2}(U_i) = Y_{it}^* - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i)$.

Additionally, we define the following notations: $m_f(u_0, \boldsymbol{Z}_{it}, \sigma) = E[f_\tau(b_{it,1}(U_i), \sigma)|U_i = u_0, \boldsymbol{Z}_{it}]$ and $m_f^*(u_0, \boldsymbol{Z}_{it,2}, \sigma) = E[f_\tau(b_{it,2}(U_i), \sigma)|U_i = u_0, \boldsymbol{Z}_{it,2}]$ with $f(\cdot)$ be one of the functions $g(\cdot)$, $g^2(\cdot)$, $\dot g(\cdot)$, and $\ddot g(\cdot)$, $m_{g_{1t}}(u_0, \boldsymbol{Z}_{i1}, \boldsymbol{Z}_{it}, \sigma) = E[g_\tau(b_{i1,1}(U_i), \sigma)\, g_\tau(b_{it,1}(U_i), \sigma)|U_i = u_0, \boldsymbol{Z}_{i1}, \boldsymbol{Z}_{it}]$, and $m_{g_{1t}}^*(u_0, \boldsymbol{Z}_{i1,2}, \boldsymbol{Z}_{it,2}, \sigma) = E[g_\tau(b_{i1,2}(U_i), \sigma) g_\tau(b_{it,2}(U_i), \sigma)|U_i = u_0, \boldsymbol{Z}_{i1,2}, \boldsymbol{Z}_{it,2}]$, respectively.

Moreover, we define several conditional variance–covariance matrices which will be used in the rest of the appendices. Firstly, let

$$\Omega_{\tau,z}(u_0, \sigma) = \tau^2 \Omega_z(u_0) - 2\tau \Omega_{zg}(u_0, \sigma) + \Omega_{zg^2}(u_0, \sigma)$$

with $\Omega_z(u_0) = E(\boldsymbol{Z}_{it}\boldsymbol{Z}_{it}'|U_i = u_0)$, $\Omega_{zg}(u_0, \sigma) = E(\boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' m_g(U_i, \boldsymbol{Z}_{it}, \sigma)|U_i = u_0)$ and $\Omega_{zg^2}(u_0, \sigma) = E(\boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' m_{g^2}(U_i, \boldsymbol{Z}_{it}, \sigma)|U_i = u_0)$. Secondly, let

$$\Omega_{\tau,z_{1t}}(u_0, \sigma) = \tau^2 \Omega_{z_{1t}}(u_0) - 2\tau \Omega_{z_{1t}g}(u_0, \sigma) + \Omega_{z_{1t}g_{1t}}(u_0, \sigma)$$

with $\Omega_{z_{1t}}(u_0) = E(\boldsymbol{Z}_{i1}\boldsymbol{Z}_{it}'|U_i = u_0)$, $\Omega_{z_{1t}g}(u_0, \sigma) = E(\boldsymbol{Z}_{i1}\boldsymbol{Z}_{it}' m_g(U_i, \boldsymbol{Z}_{it}, \sigma)|U_i = u_0)$ and $\Omega_{z_{1t}g_{1t}}(u_0, \sigma) = E(\boldsymbol{Z}_{i1}\boldsymbol{Z}_{it}' m_{g_{1t}}(U_i, \boldsymbol{Z}_{i1}, \boldsymbol{Z}_{it}, \sigma)|U_i = u_0)$. Thirdly, let

$$\Omega_{z\dot g}(u_0, \sigma) = E[\boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' m_{\dot g}(u_0, \boldsymbol{Z}_{it}, \sigma)|U_i = u_0],$$

$$\Omega_{\tau,z_2}(u_0, \sigma) = \tau^2 \Omega_{z_2}(u_0) - 2\tau \Omega_{z_2g}(u_0, \sigma) + \Omega_{z_2g^2}(u_0, \sigma)$$

with $\Omega_{z_2}(u_0) = E(\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}'|U_i = u_0)$, $\Omega_{z_2g}(u_0, \sigma) = E(\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}' m_g^*(U_i, \boldsymbol{Z}_{it,2}, \sigma)|U_i = u_0)$ and $\Omega_{z_2g^2}(u_0, \sigma) = E(\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}' m_{g^2}^*(U_i, \boldsymbol{Z}_{it,2}, \sigma)|U_i = u_0)$. Fourthly, let

$$\Omega_{\tau,z_{1t,2}}(u_0, \sigma) = \tau^2 \Omega_{z_{1t,2}}(u_0) - 2\tau \Omega_{z_{1t,2}g}(u_0, \sigma) + \Omega_{z_{1t,2}g_{1t}}(u_0, \sigma)$$

with $\Omega_{z_{1t,2}}(u_0) = E(\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}'|U_i = u_0)$, $\Omega_{z_{1t,2}g}(u_0, \sigma) = E(\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}' m_g^*(U_i, \boldsymbol{Z}_{it,2}, \sigma)|U_i = u_0)$ and $\Omega_{z_{1t,2}g_{1t}}(u_0, \sigma) = E(\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}' m_{g_{1t}}^*(U_i, \boldsymbol{Z}_{i1,2}, \boldsymbol{Z}_{it,2}, \sigma)|U_i = u_0)$. Finally, we define

$$\Omega_{z_2\dot g}(u_0, \sigma) = E[\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}' m_{\dot g}^*(U_i, \boldsymbol{Z}_{it,2}, \sigma)|U_i = u_0].$$

Thus, the asymptotic bias and variance of $\hat\gamma_\tau$ are respectively given by

$$B_{\gamma,\tau}(\sigma) = \frac{\mu_2 h_1^2}{2} e_1' E[\Omega_{z\dot g}^{-1}(U_i, \sigma)\Theta_\tau(U_i, \sigma)] \tag{23}$$

with $\Theta_\tau(U_i, \sigma) = E\{m_{\ddot g}(U_i, \boldsymbol{Z}_{it}, \sigma)\boldsymbol{Z}_{it}[\boldsymbol{Z}_{it,2}'\dot{\boldsymbol{\theta}}_\tau(U_i)]^2|U_i\}$, and

$$\Sigma_{\gamma,\tau}(\sigma) = e_1' E\{\Omega_{z\dot g}^{-1}(U_i, \sigma)[\frac{1}{T}\Omega_{\tau,z}(U_i, \sigma) + \sum_{t=2}^{T} \frac{2(T-t+1)}{T^2}\Omega_{\tau,z_{1t}}(U_i, \sigma)]\Omega_{z\dot g}^{-1}(U_i, \sigma)\}e_1. \tag{24}$$

Similarly, the asymptotic bias and variance of $\hat\beta_\tau(u_0)$ are given by

$$B_{\beta,\tau}(u_0, \sigma) \equiv B_{\beta,\tau}(u_0) = \frac{\mu_2 h_2^2}{2}\ddot{\boldsymbol{\beta}}_\tau(u_0), \tag{25}$$

and

$$\Sigma_{\beta,\tau}(u_0, \sigma) = \frac{\nu_0 e_2'}{f_U(u_0)}\Omega_{z_2\dot g}^{-1}(u_0, \sigma)[\frac{1}{T}\Omega_{\tau,z_2}(u_0, \sigma) + \sum_{t=2}^{T}\frac{2(T-t+1)}{T^2}\Omega_{\tau,z_{1t,2}}(u_0, \sigma)]\Omega_{z_2\dot g}^{-1}(u_0, \sigma)e_2. \tag{26}$$

Since $g(a, \sigma) = I_{a<0}$ when $\sigma = 0$, then $m_g(u_0, \boldsymbol{Z}_{it}, \sigma)$, $m_g^*(u_0, \boldsymbol{Z}_{it,2}, \sigma)$, $m_{g^2}(u_0, \boldsymbol{Z}_{it}, \sigma)$, $m_{g^2}^*(u_0, \boldsymbol{Z}_{it,2}, \sigma)$, $m_{g_{1t}}(u_0, \boldsymbol{Z}_{i1}, \boldsymbol{Z}_{it}, \sigma)$ and $m_{g_{1t}}^*(u_0, \boldsymbol{Z}_{i1,2}, \boldsymbol{Z}_{it,2}, \sigma)$ are all equal to the quantile level τ, and furthermore, note that $m_{\dot g}(u_0, \boldsymbol{Z}_{it}, \sigma)$, $m_{\dot g}^*(u_0, \boldsymbol{Z}_{it,2}, \sigma)$ and $m_{\ddot g}(u_0, \boldsymbol{Z}_{it}, \sigma)$ are equal to $f_{Y|U,\boldsymbol{Z}}(\boldsymbol{Z}_{it,1}'\boldsymbol{\gamma}_\tau + \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(u_0))$, $f_{Y^*|U,\boldsymbol{Z}_2}(\boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(u_0))$ and $\dot f_{Y|U,\boldsymbol{Z}}(\boldsymbol{Z}_{it,1}'\boldsymbol{\gamma}_\tau + \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(u_0))$, respectively. As a result, we can have $\Omega_{\tau,z}(u_0, \sigma) = \tau(1-\tau)\Omega_z(u_0)$, $\Omega_{\tau,z_2}(u_0, \sigma) = \tau(1-\tau)\Omega_{z_2}(u_0)$, $\Omega_{z\dot g}(u_0, \sigma) = E[\boldsymbol{Z}_{it}\boldsymbol{Z}_{it}' f_{Y|U,\boldsymbol{Z}}(\boldsymbol{Z}_{it,1}'\boldsymbol{\gamma}_\tau + \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i))|U_i = u_0] \equiv \Omega_{zf}(u_0)$, $\Omega_{z_2\dot g}(u_0, \sigma) = E[\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}' f_{Y^*|U,\boldsymbol{Z}_2}(\boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i))|U_i = u_0] \equiv \Omega_{z_2f}(u_0)$, $\Omega_{\tau,z_{1t}}(u_0, \sigma) = \tau(1-\tau)\Omega_{z_{1t}}(u_0)$, $\Omega_{\tau,z_{1t,2}}(u_0, \sigma) = \tau(1-\tau)\Omega_{z_{1t,2}}(u_0)$.

Therefore, the formula of asymptotic bias $B_{\gamma,\tau}(\sigma)$ reduces to

$$B_{\gamma,\tau} = \frac{\mu_2 h_1^2}{2} e_1' E[\Omega_{zf}^{-1}(U_i)\Theta_\tau(U_i)]$$

---

13 $\frac{\partial \ln(\lambda_\tau(a,\sigma))}{\partial a} = e^{a + \frac{(1-2\tau)\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma} + (\tau-1)\sigma\big)/[\Phi\big(\frac{a}{\sigma} - \tau\sigma\big) + e^{a + \frac{(1-2\tau)\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma} + (\tau-1)\sigma\big)]$.

with $\Theta_\tau(U_i) = E\{\boldsymbol{Z}_{it}[\boldsymbol{Z}'_{it,2}\dot{\boldsymbol{\theta}}_\tau(U_i)]^2\dot{f}_{Y|U,\boldsymbol{Z}}(\boldsymbol{Z}'_{it,1}\boldsymbol{\gamma}_\tau + \boldsymbol{Z}'_{it,2}\boldsymbol{\theta}_\tau(U_i))|U_i\}$. Similarly, the asymptotic variances $\Sigma_{\gamma,\tau}(\sigma)$ and $\Sigma_{\beta,\tau}(u_0,\sigma)$ can be respectively simplified to

$$\Sigma_{\gamma,\tau} = \tau(1-\tau)e_1'E\{\Omega_{zf}^{-1}(U_i)[\frac{1}{T}\Omega_z(U_i) + \sum_{t=2}^{T}\frac{2(T-t+1)}{T^2}\Omega_{z_{1t}}(U_i)]\Omega_{zf}^{-1}(U_i)\}e_1,$$

and

$$\Sigma_{\beta,\tau}(u_0) = \frac{\tau(1-\tau)\nu_0 e_2'}{f_U(u_0)}\Omega_{z_2 f}^{-1}(u_0)[\frac{1}{T}\Omega_{z_2}(u_0) + \sum_{t=2}^{T}\frac{2(T-t+1)}{T^2}\Omega_{z_{1t,2}}(u_0)]\Omega_{z_2 f}^{-1}(u_0)e_2.$$

Appendix C. Proof of Theorem 1

In order to establish the asymptotic theory of $\hat{\boldsymbol{\gamma}}_\tau$ in Theorem 1, the local Bahadur representation for the estimators obtained from the first stage should be derived. At first, we introduce two additional notations and definitions, $U_{ih_1} = (U_i - u_0)/h_1$ and $\psi_\tau(a,\sigma) = \tau - g_\tau(a,\sigma)$. Following Cai and Xu (2008) and Cai and Xiao (2012), we have the local Bahadur representation as

$$\sqrt{Nh_1}\begin{pmatrix}\hat{\boldsymbol{\gamma}}_\tau(u_0) - \boldsymbol{\gamma}_\tau(u_0)\\ \hat{\boldsymbol{\theta}}_{0,\tau}(u_0) - \boldsymbol{\theta}_\tau(u_0)\end{pmatrix} = \frac{\Omega_{z\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_1}Tf_U(u_0)}\sum_{j=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{jt}\psi_\tau(a_{jt,1},\sigma)K(U_{jh_1}) + o_p(1), \tag{27}$$

which is useful for establishing the asymptotic results for our estimators.

For any u_0, $\frac{\Omega_{z\dot{g}}^{-1}(u_0,\sigma)}{f_U(u_0)}\boldsymbol{Z}_{jt}\psi_\tau(a_{jt,1},\sigma)K(U_{jh_1})$ can be rewritten as

$$h_1 B^{-1}(u_0,\sigma)\{Z(u_0,\boldsymbol{Z}_{jt},\sigma) + \boldsymbol{Z}_{jt}[\psi_\tau(a_{jt,1},\sigma) - \psi_\tau(b_{jt,1}(u_0),\sigma)]K_h(U_j - u_0)\}$$

where $B(u_0,\sigma) = f_U(u_0)\Omega_{z\dot{g}}(u_0,\sigma)$, $Z(u_0,\boldsymbol{Z}_{jt},\sigma) = \boldsymbol{Z}_{jt}\psi_\tau(b_{jt,1}(u_0),\sigma)K_h(U_j - u_0)$ and $b_{jt,1}(u_0) = Y_{jt} - \boldsymbol{Z}'_{jt,1}\boldsymbol{\gamma}_\tau - \boldsymbol{Z}'_{jt,2}\boldsymbol{\theta}_\tau(u_0)$ for the jth individual with the value of smoothing variable U_j in a small neighborhood of u_0. In particular, the estimation error of $\boldsymbol{\gamma}_\tau(u_0)$ is

$$\begin{aligned}\hat{\boldsymbol{\gamma}}_\tau(u_0) - \boldsymbol{\gamma}_\tau(u_0) &\simeq \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}e_1'B^{-1}(u_0,\sigma)Z(u_0,\boldsymbol{Z}_{jt},\sigma)\\ &\quad + \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}e_1'B^{-1}(u_0,\sigma)\boldsymbol{Z}_{jt}[\psi_\tau(a_{jt,1},\sigma) - \psi_\tau(b_{jt,1}(u_0),\sigma)]K_h(U_j - u_0)\\ &\equiv \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}e_1'B^{-1}(u_0,\sigma)Z(u_0,\boldsymbol{Z}_{jt},\sigma) + B_N(u_0,\sigma),\end{aligned}$$

which holds uniformly for all u_0 under Assumption A1–A4. Thus, the estimation error of $\boldsymbol{\gamma}_\tau$ is

$$\begin{aligned}\hat{\boldsymbol{\gamma}}_\tau - \boldsymbol{\gamma}_\tau &= \frac{1}{N}\sum_{i=1}^{N}[\hat{\boldsymbol{\gamma}}_\tau(U_i) - \boldsymbol{\gamma}_\tau(U_i)]\\ &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{T}\sum_{t=1}^{T}e_1'B^{-1}(U_i,\sigma)Z(U_i,\boldsymbol{Z}_{jt},\sigma) + \frac{1}{N}\sum_{i=1}^{N}B_N(U_i,\sigma)\\ &= \frac{2}{N^2}\sum_{1\le i<j\le N}e_1'B^{-1}(U_i,\sigma)\frac{1}{T}\sum_{t=1}^{T}Z(U_i,\boldsymbol{Z}_{jt},\sigma) + \frac{1}{N}\sum_{i=1}^{N}B_N(U_i,\sigma)\\ &= \frac{1}{N^2}\sum_{1\le i<j\le N}[e_1'B^{-1}(U_i,\sigma)\frac{1}{T}\sum_{t=1}^{T}Z(U_i,\boldsymbol{Z}_{jt},\sigma) + e_1'B^{-1}(U_j,\sigma)\frac{1}{T}\sum_{t=1}^{T}Z(U_j,\boldsymbol{Z}_{it},\sigma)]\\ &\quad + \frac{1}{N}\sum_{i=1}^{N}B_N(U_i,\sigma)\\ &= \frac{N-1}{2N}\mathbb{U}_N + \mathbb{B}_N,\end{aligned}$$

where $\mathbb{B}_N = \frac{1}{N}\sum_{i=1}^{N} B_N(U_i,\sigma)$ and $\mathbb{U}_N = \frac{2}{N(N-1)}\sum_{1\le i<j\le N} p_N(\xi_i,\xi_j,\sigma)$ with

$$p_N(\xi_i,\xi_j,\sigma) = e_1' B^{-1}(U_i,\sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_i,\boldsymbol{Z}_{jt},\sigma) + e_1' B^{-1}(U_j,\sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_j,\boldsymbol{Z}_{it},\sigma),$$

and $\xi_i = (U_i,\boldsymbol{Z}_i)$ indicates all the information for i. Define $r_N(\xi_i,\sigma) = E[p_N(\xi_i,\xi_j,\sigma)|\xi_i]$, $\theta_N(\sigma) = E[r_N(\xi_i,\sigma)] = E[p_N(\xi_i,\xi_j,\sigma)]$, and $\hat{\mathbb{U}}_N = \theta_N(\sigma) + \frac{2}{N}\sum_{i=1}^{N}[r_N(\xi_i,\sigma) - \theta_N(\sigma)]$. The following two lemmas are useful to prove Theorem 1 and their detailed proofs are relegated to Appendix D.

Lemma 1. *Under the assumptions in Theorem 1, we have*

$$(i)\ \ r_N(\xi_i,\sigma) = e_1'\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\{\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_{it}\psi_\tau(b_{it,1}(U_i),\sigma)(1+o(1))\},$$

$$(ii)\ \ \theta_N(\sigma) = \mu_2 h_1^2 e_1' E[\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\Theta_\tau(U_i,\sigma)] + o(h_1^2),$$

$$(iii)\ \ Var[r_N(\xi_i,\sigma)] = \boldsymbol{\Sigma}_{\gamma,\tau}(\sigma) + o(h_1).$$

Lemma 2. *Under the assumptions in Theorem 1, we have*

$$\mathbb{B}_N = o(h_1^2).$$

Proof of Theorem 1. First, note that $E[\|p_N(\xi_i,\xi_j,\sigma)\|^2] = O(h^{-1}) = O[N(Nh_1)^{-1}] \to o(N)$ if and only if $Nh_1 \to \infty$ as $h_1 \to 0$. Lemma 3.1 in Powell et al. (1989) gives that $\sqrt{N}(\mathbb{U}_N - \hat{\mathbb{U}}_N) = o_p(1)$. Then the result follows from Lemmas 1, 2 and the Lindeberg–Lévy central limit theorem.

Appendix D. Proof of Theorem 2

The asymptotic distribution of $\hat{\boldsymbol{\beta}}_\tau(u_0)$ can be easily extracted from the asymptotic distribution of $\hat{\boldsymbol{\theta}}_{0,\tau}(u_0)$ because of $\hat{\boldsymbol{\beta}}_\tau(u_0) = e_2'\hat{\boldsymbol{\theta}}_{0,\tau}(u_0)$. Similar to the proof of Theorem 1, the local Bahadur representation for the estimators obtained from the third stage should be derived at first. First, we denote $U_{ih_2} = (U_i - u_0)/h_2$. For a given $\sqrt{N}$ consistent estimator $\hat{\boldsymbol{\gamma}}_\tau$ of $\boldsymbol{\gamma}_\tau$, we can obtain that the local Bahadur representation that is similar to (27),

$$\sqrt{Nh_2}\left(\hat{\boldsymbol{\theta}}_{0,\tau}(u_0) - \boldsymbol{\theta}_\tau(u_0)\right) = \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}Tf_U(u_0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{it,2}\psi_\tau(a_{it,2},\sigma)K(U_{ih_2}) + o_p(1),$$

which is useful to establish the asymptotic result for $\hat{\boldsymbol{\theta}}_{0,\tau}(u_0)$. The above equation can be rewritten as

$$\begin{aligned}
\sqrt{Nh_2}(\hat{\boldsymbol{\theta}}_{0,\tau} - \boldsymbol{\theta}_\tau(u_0)) &\simeq \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}Tf_U(u_0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{it,2}\psi_\tau(a_{it,2},\sigma)K(U_{ih_2})\\
&= \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}Tf_U(u_0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{Z}_{it,2}[\psi_\tau(a_{it,2},\sigma) - \psi_\tau(b_{it,2}(U_i),\sigma)]K(U_{ih_2})\\
&\quad + \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}Tf_U(u_0)}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)\\
&\equiv \boldsymbol{B}_N + \boldsymbol{\Psi}_N,
\end{aligned}$$

where $b_{it,2}(U_i) = Y_{it}^* - \boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i)$ and $Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma) = \boldsymbol{Z}_{it,2}\psi_\tau(b_{it,2}(U_i),\sigma)K(U_{ih_2})$. We will show that the first term $\boldsymbol{B}_N$ determines the asymptotic bias and the second term $\boldsymbol{\Psi}_N$ gives the asymptotic normality.

First, we work on the asymptotic normality of the second term $\boldsymbol{\Psi}_N$. Note that the first-order conditional moment of $\psi_\tau(b_{it,2}(U_i),\sigma)$ is

$$\begin{aligned}
E(\psi_\tau(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}) &= E(\tau - g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2})\\
&= \tau - E[g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]\\
&\equiv \tau - m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma).
\end{aligned}$$

The second-order conditional moments are

$$\begin{aligned}
E(\psi_\tau^2(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}) &= E[\tau^2 - 2\tau g(b_{it,2}(U_i),\sigma) + g^2(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]\\
&= \tau^2 - 2\tau E[g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}] + E[g^2(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]\\
&\equiv \tau^2 - 2\tau m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma) + m_{g^2}^*(U_i,\boldsymbol{Z}_{it,2},\sigma),
\end{aligned}$$

and

$$\begin{aligned}
& E(\psi_\tau(b_{i1,2}(U_i),\sigma)\psi_\tau(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2}) \\
&= E[\tau^2-\tau g(b_{i1,2}(U_i),\sigma)-\tau g(b_{it,2}(U_i),\sigma)+g(b_{i1,2}(U_i),\sigma)g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2}] \\
&= \tau^2-\tau E[g(b_{i1,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{i1,2}]-\tau E[g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}] \\
&\quad + E[g(b_{i1,2}(U_i),\sigma)g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2}] \\
&\equiv \tau^2-\tau m_g^*(U_i,\boldsymbol{Z}_{i1,2},\sigma)-\tau m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma)+m_{g_{1t}}^*(U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2},\sigma) \\
&= \tau^2-2\tau m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma)+m_{g_{1t}}^*(U_i,\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2},\sigma).
\end{aligned}$$

Thus,

$$\begin{aligned}
E(\boldsymbol{\Psi}_N) &= \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}f_U(u_0)}NE[Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)] \\
&= \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{\sqrt{Nh_2}f_U(u_0)}NE\{E[\boldsymbol{Z}_{it,2}\psi_\tau(b_{it,2}(U_i),\sigma)|U_i]K(U_{ih_2})\}=0
\end{aligned}$$

with $E[\boldsymbol{Z}_{it,2}\psi_\tau(b_{it,2}(U_i),\sigma)|U_i]=0$, since we know that $b_{it,2}(U_i)$ is the maximizer of the corresponding likelihood function, and

$$\begin{aligned}
& Var(\boldsymbol{\Psi}_N) \\
&= \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{h_2T^2f_U^2(u_0)}Var\{\sum_{t=1}^{T}Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)\}\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma) \\
&= \frac{\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)}{h_2T^2f_U^2(u_0)}\{TVar[Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)] \\
&\quad + \sum_{t=2}^{T}2(T-t+1)Cov(Z_2(u_0,\boldsymbol{Z}_{i1,2},\sigma),Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma))\}\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma) \\
&= \frac{\nu_0}{Tf_U(u_0)}\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma)[\Omega_{\tau,z_2}(u_0,\sigma)+\sum_{t=2}^{T}\frac{2(T-t+1)}{T}\Omega_{\tau,z_{1t,2}}(u_0,\sigma)]\Omega_{z_2\dot{g}}^{-1}(u_0,\sigma),
\end{aligned}$$

since

$$\begin{aligned}
& E[\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}'\psi_\tau^2(b_{it,2}(U_i),\sigma)K^2(U_{ih_2})] \\
&= E\{\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}'E[\psi_\tau^2(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]K^2(U_{ih_2})\} \\
&= E\{\boldsymbol{Z}_{it,2}\boldsymbol{Z}_{it,2}'[\tau^2-2\tau m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma)+m_{g^2}^*(U_i,\boldsymbol{Z}_{it,2},\sigma)]K^2(U_{ih_2})\} \\
&\equiv \nu_0h_2f_U(u_0)[\tau^2\Omega_{z_2}(u_0)-2\tau\Omega_{z_2g}(u_0,\sigma)+\Omega_{z_2g^2}(u_0,\sigma)](1+o(1)) \\
&\equiv \nu_0h_2f_U(u_0)\Omega_{\tau,z_2}(u_0,\sigma)(1+o(1)),
\end{aligned}$$

and

$$\begin{aligned}
& E[\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}'\psi_\tau(b_{i1,2}(U_i),\sigma)\psi_\tau(b_{it,2}(U_i),\sigma)K^2(U_{ih_2})] \\
&= E\{\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}'E[\psi_\tau(b_{i1,2}(U_i),\sigma)\psi_\tau(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2}]K^2(U_{ih_2})\} \\
&= E\{\boldsymbol{Z}_{i1,2}\boldsymbol{Z}_{it,2}'[\tau^2-2\tau m_g^*(U_i,\boldsymbol{Z}_{it,2},\sigma)+m_{g_{1t}}^*(U_i,\boldsymbol{Z}_{i1,2},\boldsymbol{Z}_{it,2},\sigma)]K^2(U_{ih_2})\} \\
&\equiv \nu_0h_2f_U(u_0)[\tau^2\Omega_{z_{1t,2}}(u_0)-2\tau\Omega_{z_{1t,2}g}(u_0,\sigma)+\Omega_{z_{1t,2}g_{1t}}(u_0,\sigma)](1+o(1)) \\
&\equiv \nu_0h_2f_U(u_0)\Omega_{\tau,z_{1t,2}}(u_0,\sigma)(1+o(1)).
\end{aligned}$$

Let $Q_N=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)$. Using the Cramer–Wold device, for any $\boldsymbol{d}\in R^{K_2^*}$, define $Z_{N,it}=\sqrt{\frac{h_2}{T}}\boldsymbol{d}'Z_2(u_0,\boldsymbol{Z}_{it,2},\sigma)$, then we have

$$\sqrt{Nh_2}\boldsymbol{d}'Q_N=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{N,it}=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}Z_{N,i}^*,$$

where $Z_{N,i}^*=\sum_{t=1}^{T}Z_{N,it}$, which is iid across i. Hence, it follows by the Lindeberg–Lévy central limit theorem that the asymptotic normality holds.

Next, we move to work on the first term $\boldsymbol{B}_N$. Note that

$$b_{it,2}(U_i)-a_{it,2}=\boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(u_0)+\boldsymbol{Z}_{it,2}'\dot{\boldsymbol{\theta}}_\tau(u_0)h_2U_{ih_2}-\boldsymbol{Z}_{it,2}'\boldsymbol{\theta}_\tau(U_i)=-\frac{h_2^2}{2}\boldsymbol{Z}_{it,2}'\ddot{\boldsymbol{\theta}}_\tau(\tilde{u})U_{ih_2}^2,$$

then we can get

$$g(b_{it,2}(U_i),\sigma)-g(a_{it,2},\sigma) = \dot g(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)(b_{it,2}(U_i)-a_{it,2})$$
$$= -\frac{h_2^2}{2}\dot g(a_{it,2}-\zeta\frac{h_2^2}{2}\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(\tilde u)U_{ih_2}^2,\sigma)\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(\tilde u)U_{ih_2}^2$$

and

$$\begin{aligned}
&[g(b_{it,2}(U_i),\sigma)-g(a_{it,2},\sigma)]^2\\
&= 2[g(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)-g(a_{it,2},\sigma)]\dot g(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)(b_{it,2}(U_i)-a_{it,2})\\
&= 2[\dot g(a_{it,2}+\eta(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2})-a_{it,2}),\sigma)(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2})-a_{it,2})]\\
&\quad \dot g(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)(b_{it,2}(U_i)-a_{it,2})\\
&= 2\zeta\dot g(a_{it,2}+\eta\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)\dot g(a_{it,2}+\zeta(b_{it,2}(U_i)-a_{it,2}),\sigma)(b_{it,2}(U_i)-a_{it,2})^2\\
&= O(h_2^4).
\end{aligned}$$

Thus, by definition of $\psi_\tau(a,\sigma)$, we have

$$\begin{aligned}
&E[\psi_\tau(a_{it,2},\sigma)-\psi_\tau(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]\\
&= -\frac{h_2^2}{2}E[\dot g(a_{it,2}-\zeta\frac{h_2^2}{2}\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(\tilde u)U_{ih_2}^2,\sigma)\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(\tilde u)U_{ih_2}^2|U_i,\boldsymbol{Z}_{it,2}]\\
&= -\frac{h_2^2}{2}E[\dot g(b_{it,2}(U_i),\sigma)|U_i,\boldsymbol{Z}_{it,2}]\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(U_i)U_{ih_2}^2(1+o(1))\\
&\equiv -\frac{h_2^2}{2}m^*_{\dot g}(U_i,\boldsymbol{Z}_{it,2},\sigma)\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(U_i)U_{ih_2}^2(1+o(1))
\end{aligned}$$

and

$$E\{[\psi_\tau(a_{it,2},\sigma)-\psi_\tau(b_{it,2}(U_i),\sigma)]^2|U_i,\boldsymbol{Z}_{it,2}\}=O(h_2^4).$$

Let $B_{it}=\boldsymbol{Z}_{it,2}[\psi_\tau(a_{it,2},\sigma)-\psi_\tau(b_{it,2}(U_i),\sigma)]K(U_{ih_2})$, we have

$$\begin{aligned}
E(B_{it}) &= -\frac{h_2^2}{2}E[\boldsymbol{Z}_{it,2}m^*_{\dot g}(U_i,\boldsymbol{Z}_{it,2},\sigma)\boldsymbol{Z}'_{it,2}\ddot{\boldsymbol{\theta}}_\tau(U_i)U_{ih_2}^2K(U_{ih_2})](1+o(1))\\
&= -\frac{h_2^2}{2}E\{E[\boldsymbol{Z}_{it,2}\boldsymbol{Z}'_{it,2}m^*_{\dot g}(U_i,\boldsymbol{Z}_{it,2},\sigma)|U_i]\ddot{\boldsymbol{\theta}}_\tau(U_i)U_{ih_2}^2K(U_{ih_2})\}(1+o(1))\\
&= -\frac{h_2^2}{2}E[\Omega_{z_2\dot g}(U_i,\sigma)\ddot{\boldsymbol{\theta}}_\tau(U_i)U_{ih_2}^2K(U_{ih_2})](1+o(1))\\
&= -\frac{h_2^3}{2}f_U(u_0)\mu_2\Omega_{z_2\dot g}(u_0,\sigma)\ddot{\boldsymbol{\theta}}_\tau(u_0)(1+o(1))
\end{aligned}$$

and $E[B_{it}^2]=o(h_2^4)$ and similarly, we have $E[B_{it_1}B_{it_2}]=o(h_2^4)$. Finally, we show that

$$\begin{aligned}
E(\boldsymbol{B}_N) &= \frac{\Omega^{-1}_{z_2\dot g}(u_0,\sigma)}{\sqrt{Nh_2}f_U(u_0)}NE\{\boldsymbol{Z}_{it,2}[\psi_\tau(a_{it,2},\sigma)-\psi_\tau(b_{it,2}(U_i),\sigma)]K(U_{ih_2})\}\\
&= -\frac{\Omega^{-1}_{z_2\dot g}(u_0,\sigma)}{\sqrt{Nh_2}f_U(u_0)}N\frac{h_2^3}{2}f_U(u_0)\mu_2\Omega_{z_2\dot g}(u_0,\sigma)\ddot{\boldsymbol{\theta}}_\tau(u_0)(1+o(1))\\
&= \frac{\mu_2h_2^2\sqrt{Nh_2}}{2}\ddot{\boldsymbol{\theta}}_\tau(u_0)(1+o(1)).
\end{aligned}$$

and $Var(\boldsymbol{B}_N)=o(h_2^3)$. This completes the proof of Theorem 2.

Appendix E. Proofs of lemmas

Proof of Lemma 1. For any given ξ_j and U_i, by ignoring the higher orders of Taylor expansion, one can obtain the following:

$$\begin{aligned}
E[\psi_\tau(b_{jt,1}(U_i),\sigma)|\xi_j,U_i] &= \tau-E[g(b_{jt,1}(U_i),\sigma)|\xi_j,U_i]\\
&\simeq \tau-E\{g(b_{jt,1}(U_j),\sigma)+\dot g(b_{jt,1}(U_j),\sigma)\boldsymbol{Z}'_{jt,2}[\boldsymbol{\theta}_\tau(U_j)-\boldsymbol{\theta}_\tau(U_i)]\\
&\qquad +\frac{1}{2}\ddot g(b_{jt,1}(U_j),\sigma)\{\boldsymbol{Z}'_{jt,2}[\boldsymbol{\theta}_\tau(U_j)-\boldsymbol{\theta}_\tau(U_i)]\}^2|\xi_j,U_i\}\\
&\equiv \tau-m_g(U_j,\boldsymbol{Z}_{jt},\sigma)-m_{\dot g}(U_j,\boldsymbol{Z}_{jt},\sigma)\boldsymbol{Z}'_{jt,2}[\boldsymbol{\theta}_\tau(U_j)-\boldsymbol{\theta}_\tau(U_i)]
\end{aligned}$$

$$
\begin{aligned}
&\quad -\frac{1}{2} m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\{\boldsymbol{Z}_{jt,2}'[\boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i)]\}^2 \\
&= \tau - m_g(U_j, \boldsymbol{Z}_{jt}, \sigma) - m_{\dot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}' \begin{pmatrix} 0 \\ \boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i) \end{pmatrix} \\
&\quad - \frac{1}{2} m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\{\boldsymbol{Z}_{jt,2}'[\boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i)]\}^2,
\end{aligned}
$$

since $b_{jt,1}(U_i) - b_{jt,1}(U_j) = \boldsymbol{Z}_{jt,2}'[\boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i)]$. Hence, we have

$$
\begin{aligned}
&E[Z(U_i, \boldsymbol{Z}_{jt}, \sigma)|\xi_i] \\
&= E\{\boldsymbol{Z}_{jt}\psi_\tau(b_{jt,1}(U_j), \sigma)K_h(U_j - U_i)|\xi_i\} + E\{\boldsymbol{Z}_{jt}\dot{\psi}_\tau(b_{jt,1}(U_j), \sigma)[b_{jt,1}(U_i) - b_{jt,1}(U_j)]K_h(U_j - U_i)|\xi_i\} \\
&\quad + \frac{1}{2} E\{\boldsymbol{Z}_{jt}\ddot{\psi}_\tau(b_{jt,1}(U_j), \sigma)[b_{jt,1}(U_i) - b_{jt,1}(U_j)]^2 K_h(U_j - U_i)|\xi_i\} + o(h_1^2) \\
&= -E[m_{\dot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}\boldsymbol{Z}_{jt}' \begin{pmatrix} 0 \\ \boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i) \end{pmatrix} K_h(U_j - U_i)|\xi_i] \\
&\quad - \frac{1}{2} E\{m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}\{\boldsymbol{Z}_{jt,2}'[\boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i)]\}^2 K_h(U_j - U_i)|\xi_i\} + o(h_1^2) \\
&\equiv -\mathbb{I}_1 - \frac{\mathbb{I}_2}{2} + o(h_1^2),
\end{aligned}
$$

since $E\{\boldsymbol{Z}_{jt}\psi_\tau(b_{jt,1}(U_j), \sigma)K_h(U_j - U_i)|\xi_i\} = E\{E[\boldsymbol{Z}_{jt}\psi_\tau(b_{jt,1}(U_j), \sigma)|U_j]K_h(U_j - U_i)|\xi_i\} = 0$. Then, we obtain that

$$
\begin{aligned}
\mathbb{I}_1 &= E[m_{\dot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}\boldsymbol{Z}_{jt}' \begin{pmatrix} 0 \\ \boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i) \end{pmatrix} K_h(U_j - U_i)|\xi_i] \\
&= E\{E[m_{\dot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}\boldsymbol{Z}_{jt}'|U_j] \begin{pmatrix} 0 \\ \dot{\boldsymbol{\theta}}_\tau(U_i)(U_j - U_i) + \frac{1}{2}\ddot{\boldsymbol{\theta}}_\tau(U_i)(U_j - U_i)^2 \end{pmatrix} K_h(U_j - U_i)|\xi_i\} + o(h_1^2) \\
&\equiv E[\Omega_{z\dot{g}}(U_j, \sigma) \begin{pmatrix} 0 \\ \dot{\boldsymbol{\theta}}_\tau(U_i)(U_j - U_i) + \frac{1}{2}\ddot{\boldsymbol{\theta}}_\tau(U_i)(U_j - U_i)^2 \end{pmatrix} K_h(U_j - U_i)|\xi_i] + o(h_1^2) \\
&= \int [\Omega_{z\dot{g}}(U_i, \sigma) + uh_1\dot{\Omega}_{z\dot{g}}(U_i, \sigma)] \begin{pmatrix} 0 \\ uh_1\dot{\boldsymbol{\theta}}_\tau(U_i) + \frac{(uh_1)^2}{2}\ddot{\boldsymbol{\theta}}_\tau(U_i) \end{pmatrix} K(u)[f_U(U_i) + uh_1\dot{f}_U(U_i)]du + o(h_1^2) \\
&= \mu_2 h_1^2 B(U_i, \sigma) \begin{pmatrix} 0 \\ \frac{1}{2}\ddot{\boldsymbol{\theta}}_\tau(U_i) + \frac{\dot{f}_U(U_i)}{f_U(U_i)}\dot{\boldsymbol{\theta}}_\tau(U_i) \end{pmatrix} + \mu_2 h_1^2 f_U(U_i)\dot{\Omega}_{z\dot{g}}(U_i, \sigma) \begin{pmatrix} 0 \\ \dot{\boldsymbol{\theta}}_\tau(U_i) \end{pmatrix} + o(h_1^2) \\
&= \mu_2 h_1^2 B(U_i, \sigma) \begin{pmatrix} 0 \\ \frac{1}{2}\ddot{\boldsymbol{\theta}}_\tau(U_i) + \frac{\dot{f}_U(U_i)}{f_U(U_i)}\dot{\boldsymbol{\theta}}_\tau(U_i) \end{pmatrix} - \mu_2 h_1^2 f_U(U_i)\Theta(U_i, \sigma) + o(h_1^2),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{I}_2 &= E\{m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}\{\boldsymbol{Z}_{jt,2}'[\boldsymbol{\theta}_\tau(U_j) - \boldsymbol{\theta}_\tau(U_i)]\}^2 K_h(U_j - U_i)|\xi_i\} \\
&= E\{m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}[\boldsymbol{Z}_{jt,2}'\dot{\boldsymbol{\theta}}_\tau(U_j)]^2(U_j - U_i)^2 K_h(U_j - U_i)|\xi_i\} + o(h_1^2) \\
&= E\{E\{m_{\ddot{g}}(U_j, \boldsymbol{Z}_{jt}, \sigma)\boldsymbol{Z}_{jt}[\boldsymbol{Z}_{jt,2}'\dot{\boldsymbol{\theta}}_\tau(U_j)]^2|U_j\}(U_j - U_i)^2 K_h(U_j - U_i)|\xi_i\} + o(h_1^2) \\
&\equiv E[\Theta(U_j, \sigma)(U_j - U_i)^2 K_h(U_j - U_i)|\xi_i] + o(h_1^2) \\
&= \mu_2 h_1^2 f_U(U_i)\Theta_\tau(U_i, \sigma) + o(h_1^2).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
r_N(\xi_i, \sigma) &= E[e_1' B^{-1}(U_i, \sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_i, \boldsymbol{Z}_{jt}, \sigma) + e_1' B^{-1}(U_j, \sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_j, \boldsymbol{Z}_{it}, \sigma)|\xi_i] \\
&= E[e_1' B^{-1}(U_i, \sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_i, \boldsymbol{Z}_{jt}, \sigma)|\xi_i] + E[e_1' B^{-1}(U_j, \sigma)\frac{1}{T}\sum_{t=1}^{T} Z(U_j, \boldsymbol{Z}_{it}, \sigma)|\xi_i] \\
&= o(h_1) + \frac{1}{T}\sum_{t=1}^{T} e_1' \int B^{-1}(u, \sigma)\boldsymbol{Z}_{it}\psi_\tau(b_{it,1}(u), \sigma)K_h(U_i - u)f_U(u)du
\end{aligned}
$$

$$= \frac{1}{T}\sum_{t=1}^{T} e_1' B^{-1}(U_i,\sigma)\boldsymbol{Z}_{it}\psi_\tau(b_{it,1}(U_i),\sigma)f_U(U_i)(1+o(1))$$

$$= e_1'\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\{\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_{it}\psi_\tau(b_{it,1}(U_i),\sigma)(1+o(1))\}$$

and furthermore, we obtain that

$$\begin{aligned}\theta_N(\sigma) &= E[p_N(\xi_i,\xi_j,\sigma)] = 2E[e_1'B^{-1}(U_i,\sigma)Z(U_i,\boldsymbol{Z}_{jt},\sigma)]\\ &= -\mu_2 h_1^2\{E[e_1'\begin{pmatrix} 0 \\ \ddot{\boldsymbol{\theta}}_\tau(U_i)+2\frac{\dot{f}_U(U_i)}{f_U(U_i)}\dot{\boldsymbol{\theta}}_\tau(U_i)\end{pmatrix}] - E[e_1'\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\Theta_\tau(U_i,\sigma)]\} + o(h_1^2)\\ &= \mu_2 h_1^2 e_1' E[\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\Theta_\tau(U_i,\sigma)] + o(h_1^2).\end{aligned}$$

Similar to the derivation of $Var(\boldsymbol{\Psi}_N)$, it follows that

$$\begin{aligned}Var[r_N(\xi_i)] &= e_1'E\{\Omega_{z\dot{g}}^{-1}(U_i,\sigma)[\frac{1}{T}\Omega_{\tau,z}(U_i,\sigma)\\ &\quad+\sum_{t=2}^{T}\frac{2(T-t+1)}{T^2}\Omega_{\tau,z_{1t}}(U_i,\sigma)]\Omega_{z\dot{g}}^{-1}(U_i,\sigma)\}e_1 + o(h_1)\\ &\equiv \boldsymbol{\Sigma}_\gamma(\sigma)+o(h_1).\end{aligned}$$

Therefore, Lemma 1 is established.

Proof of Lemma 2. Similar to the proof of Theorem 2, we can show that

$$E(\mathbb{B}_N) = o(h_1^2).$$

The lemma is established due to the fact that $Var(\mathbb{B}_N) = O(\frac{1}{N}h_1^2)$.

Appendix F

Derivation of (9)

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} exp[-\rho_\tau(a-v)-\frac{v^2}{2\sigma^2}]dv\\ &= \frac{1}{\sqrt{2\pi}\sigma}\{\int_{-\infty}^{a} exp[-\tau(a-v)-\frac{v^2}{2\sigma^2}]dv + \int_{a}^{\infty} exp[-(\tau-1)(a-v)-\frac{v^2}{2\sigma^2}]dv\}\\ &= \frac{1}{\sqrt{2\pi}\sigma}\{e^{-\tau a}\int_{-\infty}^{a} exp[-\frac{1}{2\sigma^2}(v^2-2\tau\sigma^2 v)]dv + e^{-(\tau-1)a}\int_{a}^{\infty} exp[-\frac{1}{2\sigma^2}(v^2-2(\tau-1)\sigma^2 v)]dv\}\\ &= e^{-\tau a+\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big) + e^{-(\tau-1)a+\frac{(\tau-1)^2\sigma^2}{2}}[1-\Phi\big(\frac{a}{\sigma}-(\tau-1)\sigma\big)]\\ &= e^{-\tau a+\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big) + e^{-(\tau-1)a+\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)\\ &= \begin{cases} e^{-\tau a}[e^{\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{a+\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)] & \text{, when } a\geq 0\\ e^{-(\tau-1)a}[e^{-a+\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)] & \text{, when } a<0\end{cases}\\ &= e^{-\rho_\tau(a)}I_{a\geq 0}[e^{\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{a+\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)]\\ &\quad+e^{-\rho_\tau(a)}I_{a<0}[e^{-a+\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)]\\ &= e^{-\rho_\tau(a)}[e^{\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)e^a](I_{a\geq 0}+e^{-a}I_{a<0})\\ &\equiv e^{-\rho_\tau(a)}\lambda_\tau(a,\sigma)(I_{a\geq 0}+e^{-a}I_{a<0})\end{aligned}$$

Therefore, Eq. (9) is established.

Derivation of (12)

Differentiation $\max_\sigma \sum_{i=1}^{N}\sum_{t=1}^{T}\ln(\lambda_\tau(a_{it},\sigma))$ with respect to σ leads to

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\frac{\partial\lambda_\tau(a_{it},\sigma)}{\partial\sigma}}{\lambda_\tau(a_{it},\sigma)}|_{\sigma=\tilde{\sigma}}=0.$$

By definition of $\lambda_\tau(a,\sigma)=e^{\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+e^{\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)e^{a}$,

$$\begin{aligned}\frac{\partial\lambda_\tau(a,\sigma)}{\partial\sigma} &= \tau^2\sigma e^{\frac{\tau^2\sigma^2}{2}}\Phi\big(\frac{a}{\sigma}-\tau\sigma\big)+(\tau-1)^2\sigma e^{a+\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)\\ &\quad+(-\frac{a}{\sigma^2}-\tau)e^{\frac{\tau^2\sigma^2}{2}}\phi\big(\frac{a}{\sigma}-\tau\sigma\big)+(\frac{a}{\sigma^2}+\tau-1)e^{a+\frac{(\tau-1)^2\sigma^2}{2}}\phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)\\ &= \tau^2\sigma\lambda_\tau(a,\sigma)+(1-2\tau)\sigma e^{a+\frac{(\tau-1)^2\sigma^2}{2}}\Phi\big(-\frac{a}{\sigma}+(\tau-1)\sigma\big)-e^{\frac{\tau^2\sigma^2}{2}}\phi\big(\frac{a}{\sigma}-\tau\sigma\big).\end{aligned}$$

Thus,

$$NT\tau^2\tilde{\sigma}+\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{-\phi\big(\frac{a_{it}}{\tilde{\sigma}}-\tau\tilde{\sigma}\big)+(1-2\tau)\tilde{\sigma}e^{a_{it}+\frac{(1-2\tau)\tilde{\sigma}^2}{2}}\Phi\big(-\frac{a_{it}}{\tilde{\sigma}}+(\tau-1)\tilde{\sigma}\big)}{\Phi\big(\frac{a_{it}}{\tilde{\sigma}}-\tau\tilde{\sigma}\big)+e^{a_{it}+\frac{(1-2\tau)\tilde{\sigma}^2}{2}}\Phi\big(-\frac{a_{it}}{\tilde{\sigma}}+(\tau-1)\tilde{\sigma}\big)}=0.$$

Therefore, Eq. (12) is established.

References

Abrevaya, J., Dahl, C., 2008. The effects of birth inputs on birth weight: Evidence from quantile estimation on panel data. J. Bus. Econom. Statist. 26, 379–397.

Aitken, B., Harrison, A., 1999. Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. Amer. Econ. Rev. 89, 605–618.

Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S., 2004. FDI and economic growth: The role of local financial markets. J. Int. Econ. 64, 89–112.

Angrist, J., Chernozhukov, V., Fernandez-Val, I., 2006. Quantile regression under misspecification with an application to the U.S. wage structure. Econometrica 74, 539–563.

Blomstrom, M., Lipsey, R., Zejan, M., 1992. What Explains Developing Country Growth? Working Paper No. 4132, NBER.

Blomstrom, M., Persson, H., 1983. Foreign investment and spillover efficiency in an underdeveloped economy: evidence from the Mexican manufacturing industry. World Dev. 11, 493–501.

Borensztein, E., De Gregorio, J., Lee, J.-W., 1998. How does foreign direct investment affect economic growth? J. Int. Econ. 45, 115–135.

Cai, Z., 2010. Functional coefficient models for economic and financial data. In: Ferraty, Y., Romain, F. (Eds.), Handbook of Functional Data Analysis. Oxford University Press, Oxford, pp. 166–186.

Cai, Z., Das, M., Xiong, H., Wu, X., 2006. Functional coefficient instrumental variables models. J. Econometrics 133, 207–241.

Cai, Z., Masry, E., 2000. Nonparametric estimation of additive nonlinear ARX time series: Local linear fitting and projection. Econometric Theory 16, 465–501.

Cai, Z., Xiao, Z., 2012. Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. J. Econometrics 167, 413–425.

Cai, Z., Xu, X., 2008. Nonparametric quantile estimations for dynamic smooth coefficient models. J. Amer. Statist. Assoc. 103, 1595–1608.

Canay, I., 2011. A simple approach to quantile regression for panel data. Econom. J. 14, 368–386.

Carkovic, M., Levine, R., 2005. Does foreign direct investment accelerate economic growth?. In: Moran, T., Graham, E., Blomstrom, M. (Eds.), Does Foreign Investment Promote Development? Peterson Institute Press, Washington.

Chamberlain, G., 1982. Multivariate regression models for panel data. J. Econometrics 18, 5–46.

Chamberlain, G., 1984. Panel data. In: Griliches, Z., Intriligator, M. (Eds.), Handbook of Econometrics, vol. II. North-Holland, Amsterdam, pp. 1247–1318.

Chaudhuri, P., 1991. Global nonparametric estimation of conditional quantile functions and their derivatives. J. Multivariate Anal. 39, 246–269.

Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. J. Econometrics 132, 491–525.

De Gooijer, J.G., Zerom, D., 2003. On additive conditional quantiles with high-dimensional covariates. J. Amer. Statist. Assoc. 98, 135–146.

De Gregorio, J., 1992. Economic growth in Latin America. J. Dev. Econ. 39, 58–84.

De Mello, R., 1999. FDI-led growth: Evidence from time series and panel data. Oxf. Econ. Pap. 51, 133–151.

Durham, J.B., 2004. Absorptive capacity and the effects of foreign direct investment and equity foreign portfolio investment on economic growth. Eur. Econ. Rev. 48, 285–306.

Durlauf, S.N., Kourtellos, A., Tan, C., 2008. Are any growth theories robust? Econ. J. 118, 329–346.

Fan, J., Yao, Q., Cai, Z., 2003. Adaptive varying-coefficient linear models. J. R. Stat. Soc. Ser. B Stat. Methodol. 65, 57–80.

Galvao, A., 2011. Quantile regression for dynamic panel data with fixed effects. J. Econometrics 164, 142–157.

Gamper-Rabindran, S., Khan, S., Timmins, C., 2010. The impact of piped water provision on infant mortality in Brazil: a quantile panel data approach. J. Dev. Econ. 92, 188–200.

Ghosh, M., Wang, W., 2009. Does FDI accelerate economic growth? The OECD experience based on panel data estimation for the period 1980–2004. Global Econ. J. 9, 1–21.

Grog, H., Strobl, E., 2001. Multinational companies and productivity spillovers: an meta-analysis. Econ. J. 111, 723–739.

Haddad, M., Harrison, A., 1993. Are there positive spillovers from direct foreign investment? Evidence from panel data for Morocco. J. Dev. Econ. 42, 51–74.

Hansen, B., 2000. Sample splitting and threshold estimation. Econometrica 68, 575–603.

Hausman, J., Lu, Y., Palmer, C., 2014. Errors in Dependent Variable of Quantile Regression Model. Working Paper. Department of Economics, MIT.

He, X., Liang, H., 2000. Quantile regression estimates for a class of linear and partially linear errors-in-variable models. Statist. Sinica 10, 129–140.

He, X., Ng, P., 1999. Quantile splines with several covariates. J. Statist. Plann. Inference 75, 343–352.

He, X., Ng, P., Portnoy, S., 1998. Bivariate quantile smoothing splines. J. R. Stat. Soc. Ser. B Stat. Methodol. 60, 537–550.

He, X., Portnoy, S., 2000. Some asymptotic results on bivariate quantile splines. J. Statist. Plann. Inference 91, 341–349.

Henderson, D., Papageorgiou, C., Parmeter, C., 2011. Growth empirics without parameters. Econ. J. 122, 125–154.

Hermes, N., Lensink, R., 2003. Foreign direct investment, financial development and economic growth. J. Dev. Stud. 40, 142–163.
Honda, T., 2004. Quantile regression in varying coefficient models. J. Statist. Plann. Inference 121, 113–125.
Horowitz, J., Lee, S., 2005. Nonparametric estimation of an additive quantile regression model. J. Amer. Statist. Assoc. 100, 1238–1249.
Koenker, R., 2004. Quantile regression for longitudinal data. J. Multivariate Anal. 91, 74–89.
Koenker, R., 2005. Quantile Regression. Cambridge University Press, New York.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46, 33–50.
Koenker, R., Bassett, G., 1982. Robust tests for heteroscedasticity based on regression quantiles. Econometrica 50, 43–61.
Komunjer, I., 2005. Quasi-maximum likelihood estimation for conditional quantiles. J. Econometrics 128, 137–164.
Kottaridi, C., Stengos, T., 2010. Foreign direct investment, human capital and non-linearities in economic growth. J. Macroeconom. 32, 858–871.
Lamarche, C., 2010. Robust penalized quantile regression estimation for panel data. J. Econometrics 157, 396–408.
Lee, S., 2003. Efficient semiparametric estimation of partially linear quantile regression model. Econometric Theory 19, 1–31.
Li, D., Li, Q., Li, Z., 2015. Nonparametric Estimation of Conditional Quantile Regression with Mixed Discrete and Continuous Data. Working Paper. Department of Economics, Texas A&M University.
Lipsey, R., 2003. Foreign direct investment, growth, and competitiveness in developing countries. In: Cornelius, P. (Ed.), The Global Competitiveness Report 2002–2003. Oxford University Press, New York.
Luiz, R., De Mello, J., 1999. Foreign direct investment-led growth: Evidence from time series and panel data. Oxf. Econ. Pap. 51, 133–151.
Maasoumi, E.J., Racine, J.S., Stengos, T., 2007. Growth and convergence: A profile of distribution dynamics and mobility. J. Econometrics 136, 483–508.
Nunnenkamp, P., 2004. To what extent can foreign direct investment help achieve international development goals? World Econ. 27, 657–677.
Powell, J., Stock, J., Stoker, T., 1989. Semiparametric estimation of index coefficients. Econometrica 57, 1403–1430.
Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56, 931–954.
Speckman, P., 1988. Kernel smoothing partial linear models. J. R. Stat. Soc. Ser. B Stat. Methodol. 50, 413–426.
Yu, K., Lu, Z., 2004. Local linear additive quantile regression. Scand. J. Stat. 31, 333–346.