

A panel quantile model via correlated random effects approach for testing pecking order theory

Zongwu Cai^a, Meng Shi^b, Wuqing Wu^c, and Yue Zhao^{c,d}

^aDepartment of Economics, University of Kansas, Lawrence, Kansas, USA; ^bAcademy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China; ^cSchool of Business, Renmin University of China, Beijing, China; ^dSchool of Accountancy, Singapore Management University, Singapore, Singapore

ABSTRACT

This article investigates the relative importance of internal and external sources of funds in financing activities across different levels of investment activities by proposing a panel data quantile regression model with correlated random effects, accounting for heteroscedasticity in both firm-specific individuals and distribution of investment. A new estimation method, which takes the influence of disturbance into account, is proposed by using the integrated quasi-likelihood function for the conditional quantile model and Laplace approximation. The large sample theory for the proposed estimator and the corresponding asymptotic χ^2 test are investigated. A Monte Carlo simulation is conducted to examine the finite sample performance of the proposed estimator. Finally, empirical results find strong evidence that the financing hierarchy of U.S. firms is in accordance with the first rung of the pecking order theory across all levels of investments from 10% to 90%, but for the second rung of the pecking order theory, only at 30% to 70% levels of investments.

ARTICLE HISTORY

Received 18 June 2025
Accepted 5 September 2025

KEYWORDS

Correlated random effects; panel data; pecking order theory; quantile regression model; quasi-maximum likelihood estimator

JEL CLASSIFICATION:

C33; C31

1. Introduction

The pecking order theory in finance is a widely held belief that firms prefer internal (retained earnings) to external funds, and prefer debt to equity if internal funds are exhausted and external funds are sought (Myers, 1984; Myers and Majluf, 1984). For more than 30 years, a substantial empirical literature has examined whether and when there is a departure of observed corporate financing behaviors from the pecking order theory in practice (Bharath, Pasquariello, and Wu, 2009; Brennan and Kraus, 1987; Brown et al., 2019; Brown, Fazzari, and Peterson, 2009; Chay et al., 2015; Fama and French, 2005; Frank and Goyal, 2003; Jung, Kim, and Stulz, 1996; Shyam-Sunder and Myers, 1999). Unfortunately, no consensus has been reached.

Indeed, most of these studies employ a mean model, of which the results and conclusions are not warranted for two main reasons. On the one hand, the traditional mean regression (e.g., OLS) assumes an unchanging association at different levels of investment scales. However, this article predicts that the relationship between financing sources and investments varies across different investment scales. On the other hand, the estimation of linear model or fixed effect model is unbiased only when the dependent variable is normally distributed or with a central tendency. In contrast, the distribution of investment amount is usually highly skewed and heavily tailed, see Figure 5 later for investment amount in year 2016, which shows clearly that the distribution of investment amount is highly right-tailed skewed. Thus, the mean regression estimates are likely to be biased and probably not truly reflect the behavior of data in tail

regions. To the best of our knowledge, the article by Chay et al. (2015) was among the first and the latest study to investigate the responsiveness of investments to internal funds, debt issues and equity issues taking a quantile regression based on panel data. However, limited by the difficulties in the theoretical modeling and computational challenges to control firm and year fixed effects, their findings may suffer from endogeneity problems while this article takes a further step to introduce a linear quantile model for panel data with fixed effects.

Recently, there is growing literature to study panel quantile models with individual (fixed) effects. The main reason is that the commonly used method to eliminate unobservable individual effects in the conditional mean model, taking difference, is not available in the conditional quantile regression model. Among this literature, Koenker (2004) first introduced a location-shift quantile regression model for panel data with fixed effects which assumes that the fixed effect is purely a location shift and time period T goes to infinity. The fixed effects are viewed as nuisance parameters and regularized by L_1 -penalty to shrinkage them to common values. Furthermore, Lamarche (2010) discussed the choice of the tuning parameter λ in Koenker (2004), and proposed an optimal selection method. Galvao, Lamarche, and Lima (2013) made an extension to a class of censored quantile regression models for panel data with fixed effects by applying fixed effects quantile regression to subsets of observations selected either parametrically or non parametrically. For more on the most recent work, the reader is referred to the papers by Gu and Volgushev (2019), Machado and Santos Silva (2019), Powell (2022), and the references therein. However, most of the aforementioned studies rely on the critical assumption that the time period T is large or goes to infinity.

Alternative method of estimating individual effects in the panel quantile model is to view individual effects as correlated random effects when T is small or finite. In the framework of Chamberlain (1982, 1984), Abrevaya and Dahl (2008) employed a correlated random-effects quantile model to estimate the effects of mother smoking on the entire birthweight distribution. The correlated random effects are assumed to be a linear projection on observable independent variables plus a disturbance, while Harding and Lamarche (2017) relaxed the assumption in Abrevaya and Dahl (2008), by allowing the unobserved α_i to be arbitrarily related to observable variables, to increase the flexibility in the model specification.

Motivated by controlling the firm-specific heterogeneity in testing the pecking order theory under the quantile framework, we propose a panel data quantile regression model with individual effects, which is the usual model specification in conditional mean model for panel data. We view the individual effects as the correlated random effects in the framework of Chamberlain (1982) and Abrevaya and Dahl (2008), where we do not require the time period T going to infinity, as is the main assumption in many models (Galvao, Lamarche, and Lima, 2013; Gu and Volgushev, 2019; Koenker, 2004, Lamarche, 2010; Machado and Santos Silva, 2019; Powell, 2022). Actually, $T \geq 2$ is only needed for model identification.

The main contributions of this article, compared with the previous panel quantile model with correlated random effects, focus on the following three points. First, we imposes a technique, which integrates out the unobservable disturbance, to obtain the integrated quasi-likelihood function. The new estimation method takes the influence of disturbance into account, rather than choosing to *ignore* it as in Abrevaya and Dahl (2008). Second, an asymptotic χ^2 test is established to obtain the theoretical p -value and confidence interval for the parameters of interest, instead of using the Bootstrap estimates like much literature, e.g., Harding and Lamarche (2017). Third, we apply the Laplace approximation to relax the normality assumption on disturbance, which is employed in Cai, Chen, and Fang (2018). As an application of the proposed method, our empirical study finds a strong evidence that the financing hierarchy of U.S. firms is in accordance with the first rung of the pecking order theory across all levels of investments from 10% to 90%.

The remainder of this article is organized as follows. Section 2 introduces the proposed quantile regression model for panel data with correlated random effects. The estimation method and asymptotic properties of our estimator are also provided. Section 3 examines the finite sample performance of the proposed estimator, compared with the Abrevaya and Dahl (2008) estimator, via Monte Carlo simulations, and the results are consistent with the asymptotic theories in Section 2. In Section 4, we apply the model to testing the pecking order theory using non financial and non utility U.S. firms from

Compustat database. Section 5 offers conclusions and suggests some limitations and possible extensions for future research. All technical proofs are relegated to Appendix together with the list of assumptions needed for the asymptotic properties for the proposed estimator.

2. Econometric modeling

2.1. Model setup

We first introduce the econometric model studied in this article. Given $\tau \in (0, 1)$, we consider the following *panel data quantile* regression model with *correlated random effects* (PDQ-CRE). Let y_{it} , a scalar dependent variable, be the observation on i th individual at time t for $1 \leq i \leq N$ and $1 \leq t \leq T$. The conditional quantile model is given by

$$q_\tau(y_{it} | \mathbf{x}_{it}, \alpha_{i,\tau}) = \boldsymbol{\delta}_\tau^T \mathbf{x}_{it} + \alpha_{i,\tau}, \tag{1}$$

where $q_\tau(y_{it} | \mathbf{x}_{it}, \alpha_{i,\tau})$ is the τ th quantile of y_{it} given \mathbf{x}_{it} and $\alpha_{i,\tau}$, \mathbf{x}_{it} is regressor with $K \times 1$ dimensions, $\boldsymbol{\delta}_\tau$ denotes a vector of constant parameters, and $\alpha_{i,\tau}$ is an individual effect dependent on τ . In this model, we assume both parameters $\boldsymbol{\delta}_\tau$ and individual effect $\alpha_{i,\tau}$ can be dependent on τ . When T goes to infinity, one can use estimation method studied by Koenker (2004). But in this article, we consider the case that T is a fixed finite number or small. Thus, we follow the idea in Chamberlain (1982, 1984), similar to Abrevaya and Dahl (2008) and Harding and Lamarche (2017), and view the individual effect $\alpha_{i,\tau}$ as a correlated random effect which is allowed to be correlated with covariates $\mathbf{x}_i = (\mathbf{x}_{i1}^T, \dots, \mathbf{x}_{iT}^T)^T$; that is,

$$\alpha_{i,\tau} = \alpha_\tau(\mathbf{x}_i) + v_{i,\tau},$$

where $\alpha_\tau(\cdot)$ is an unknown function of \mathbf{x}_i , and $v_{i,\tau}$ is an unobserved disturbance. The unknown function $\alpha_\tau(\mathbf{x}_i)$ is approximated by a linear projection on \mathbf{x}_i such that

$$\alpha_\tau(\mathbf{x}_i) = \sum_{s=1}^T \boldsymbol{\eta}_{s,\tau}^T \mathbf{x}_{is} + \eta_{0,\tau},$$

where $\eta_{0,\tau}$ is an unknown constant and $\{\boldsymbol{\eta}_{i,\tau}\}_{i=1}^T$ is a sequence of $K \times 1$ vector of unknown constant parameters or nuisance parameters. Therefore, model (1) can be rewritten as

$$q_\tau(y_{it} | \mathbf{x}_i, v_{i,\tau}) = \boldsymbol{\delta}_\tau^T \mathbf{x}_{it} + \sum_{s=1}^T \boldsymbol{\eta}_{s,\tau}^T \mathbf{x}_{is} + \eta_{0,\tau} + v_{i,\tau}. \tag{2}$$

Although the linear approximation may seem to be restrictive, it is the usual specification employed by empirical researchers in cross sectional applications and can be seen as reduced-form approximation to the true conditional quantile, as discussed in Abrevaya and Dahl (2008). Furthermore, similar to Cai, Chen, and Fang (2018), a semiparametric form of $\alpha(\mathbf{x}_i)$ might be considered. Also, it is interesting to notice that model (2) happens to be a special case of the semiparametric panel data model proposed in Cai, Chen, and Fang (2018) if the non parametric part of the model is not considered. Finally, model (2) covers the following conditional quantile regression model with error in the dependent variable for cross-sectional data as a special case when $T = 1$,

$$q_\tau(y_i | \mathbf{x}_i, v_i) = q_\tau(\mathbf{x}_i) + v_i,$$

where $q_\tau(\mathbf{x}_i)$ is the τ th quantile of true y_i^* with $y_i = y_i^* + v_i$ and v_i is the measurement error. Indeed, Hausman et al. (2021) considered the case that $q_\tau(\mathbf{x}_i)$ is a linear function of \mathbf{x}_i .

2.2. Estimation procedures

2.2.1. Pooled regression method

We consider the following equations:

$$q_\tau(y_{it}|\mathbf{x}_i, v_{i,\tau}) = \boldsymbol{\delta}_\tau^\top \mathbf{x}_{it} + \sum_{s=1}^T \boldsymbol{\eta}_{s,\tau}^\top \mathbf{x}_{is} + \eta_{0,\tau} + v_{i,\tau} = (\boldsymbol{\delta}_\tau + \boldsymbol{\eta}_{t,\tau})^\top \mathbf{x}_{it} + \sum_{\substack{s=1 \\ s \neq t}}^T \boldsymbol{\eta}_{s,\tau}^\top \mathbf{x}_{is} + \eta_{0,\tau} + v_{i,\tau},$$

and

$$q_\tau(y_{ir}|\mathbf{x}_i, v_{i,\tau}) = \boldsymbol{\delta}_\tau^\top \mathbf{x}_{ir} + \sum_{s=1}^T \boldsymbol{\eta}_{s,\tau}^\top \mathbf{x}_{is} + \eta_{0,\tau} + v_{i,\tau} = (\boldsymbol{\delta}_\tau + \boldsymbol{\eta}_{r,\tau})^\top \mathbf{x}_{ir} + \sum_{\substack{s=1 \\ s \neq r}}^T \boldsymbol{\eta}_{s,\tau}^\top \mathbf{x}_{is} + \eta_{0,\tau} + v_{i,\tau},$$

where $t \neq r$. Obviously, the effects of \mathbf{x}_{it} upon the conditional quantile $q_\tau(y_{it}|\mathbf{x}_i, v_{i,\tau})$ are through two channels: a direct effect $\boldsymbol{\delta}_\tau^\top \mathbf{x}_{it}$ and an indirect effect $\boldsymbol{\eta}_{t,\tau}^\top \mathbf{x}_{it}$, a part of the unobservable $\alpha_{i,\tau}$. Hence, $\boldsymbol{\delta}_\tau$ can be identified by the following equation:

$$\boldsymbol{\delta}_\tau = \frac{\partial q_\tau(y_{it}|\mathbf{x}_i, v_{i,\tau})}{\partial \mathbf{x}_{it}} - \frac{\partial q_\tau(y_{ir}|\mathbf{x}_i, v_{i,\tau})}{\partial \mathbf{x}_{it}} = \frac{\partial q_\tau(y_{ir}|\mathbf{x}_i, v_{i,\tau})}{\partial \mathbf{x}_{ir}} - \frac{\partial q_\tau(y_{it}|\mathbf{x}_i, v_{i,\tau})}{\partial \mathbf{x}_{ir}}.$$

That is, $\boldsymbol{\delta}_\tau$ tells us how much \mathbf{x}_{it} affects $q_\tau(y_{it}|\mathbf{x}_i, v_{i,\tau})$ through the unobservable $\alpha_{i,\tau}$.

We adopt the pooled regression as in Abrevaya and Dahl (2008) by stacking \mathbf{x}_{it} at different time t together. Let

$$\begin{pmatrix} y_{11} \\ \vdots \\ y_{1T} \\ \vdots \\ y_{i1} \\ \vdots \\ y_{iT} \\ \vdots \\ y_{N1} \\ \vdots \\ y_{NT} \end{pmatrix} \text{ .and} \begin{pmatrix} 1 & \mathbf{x}_{11}^\top & \mathbf{x}_{11}^\top & \cdots & \mathbf{x}_{1T}^\top \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & \mathbf{x}_{1T}^\top & \mathbf{x}_{11}^\top & \cdots & \mathbf{x}_{1T}^\top \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & \mathbf{x}_{i1}^\top & \mathbf{x}_{i1}^\top & \cdots & \mathbf{x}_{iT}^\top \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & \mathbf{x}_{iT}^\top & \mathbf{x}_{i1}^\top & \cdots & \mathbf{x}_{iT}^\top \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & \mathbf{x}_{N1}^\top & \mathbf{x}_{N1}^\top & \cdots & \mathbf{x}_{NT}^\top \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & \mathbf{x}_{NT}^\top & \mathbf{x}_{N1}^\top & \cdots & \mathbf{x}_{NT}^\top \end{pmatrix} \equiv \begin{pmatrix} \mathbf{g}_{11}^\top \\ \vdots \\ \mathbf{g}_{1T}^\top \\ \vdots \\ \mathbf{g}_{i1}^\top \\ \vdots \\ \mathbf{g}_{iT}^\top \\ \vdots \\ \mathbf{g}_{N1}^\top \\ \vdots \\ \mathbf{g}_{NT}^\top \end{pmatrix} \equiv \mathbf{G}.$$

be the dependent and the explanatory variables of the quantile regression, respectively, where \mathbf{g}_{11}^\top denotes the first row vector in the above design matrix \mathbf{G} , \mathbf{g}_{12}^\top denotes the second row vector, etc. Then, we now consider the following transformed model from (2),

$$q_\tau(y_{it}|\mathbf{g}_{it}, v_{i,\tau}) = \boldsymbol{\xi}_\tau^\top \mathbf{g}_{it} + v_{i,\tau}, \tag{3}$$

where $\boldsymbol{\xi}_\tau$ includes $\boldsymbol{\delta}_\tau$ and $\boldsymbol{\eta}_\tau$'s.

2.2.2. Quasi-likelihood approach

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

$$\hat{\boldsymbol{\xi}}_\tau = \arg \min_{\boldsymbol{\xi}_\tau} \sum_{t=1}^T \rho_\tau(y_t - q_\tau(y_t|\mathbf{x}_t; \boldsymbol{\xi}_\tau)),$$

where $\rho_\tau(x) = x(\tau - I_{x < 0})$ is the check function, $q_\tau(y_t | \mathbf{x}_t; \xi_\tau)$ is the conditional quantile regression function of y_t given \mathbf{x}_t , and satisfies $P(y_t \leq q_\tau(y_t | \mathbf{x}_t; \xi_\tau)) = \tau$. Komunjer (2005) applied a different optimization approach by proposing a class of quasi-maximum likelihood estimators (QMLE), obtained by solving

$$\hat{\xi}_\tau = \arg \max_{\xi_\tau} \sum_{t=1}^T \ln l_t(y_t, q_\tau(y_t | \mathbf{x}_t; \xi_\tau)),$$

where $l_t(\cdot)$ is the conditional quasi-likelihood function y_t given \mathbf{x}_t at time t .

Using the aforementioned idea, we consider a class of QMLEs for the conditional quantile regression model defined in (2), obtained by solving the maximization of a quasi-likelihood function for the τ th conditional quantile

$$\hat{\xi}_\tau = \arg \max_{\xi_\tau} \sum_{i=1}^N \sum_{t=1}^T \ln l(y_{it}, q_\tau(y_{it} | \mathbf{g}_{it}, v_{i,\tau}; \xi_\tau)),$$

where $l(y_{it}, q_\tau(y_{it}, \mathbf{g}_{it}, v_{i,\tau}; \xi_\tau))$ is the quasi-likelihood function for the τ th conditional quantile on individual i at time t . Indeed, model (2) can be expressed as follows:

$$q_\tau(y_{it} | \mathbf{x}_i, v_{i,\tau}) = q_{\tau,-v}(y_{it} | \mathbf{x}_i) + v_{i,\tau},$$

where $q_{\tau,-v}(y_{it} | \mathbf{x}_i) = \delta_\tau^\top \mathbf{x}_{it} + \sum_{s=1}^T \boldsymbol{\eta}_{s,\tau}^\top \mathbf{x}_{is} + \eta_{0,\tau}$. Hence, $v_{i,\tau}$ in the quasi-likelihood function $l(y_{it}, q_\tau(y_{it} | \mathbf{x}_i))$ can be integrated out so that we obtain the integrated quasi-likelihood function

$$l(y_{it}, q_{\tau,-v}(y_{it} | \mathbf{x}_i)) = \int l(y_{it}, q_\tau(y_{it} | \mathbf{x}_i, v_{i,\tau})) f(v_{i,\tau}) dv_{i,\tau}, \tag{4}$$

by assuming $\{v_{i,\tau}\}$ are independent and identically distributed (i.i.d.) and independent of $\{\mathbf{x}_{it}\}$, where $f(v)$ is the density function of $v_{i,\tau}$, which is assumed to be free of parameter τ .

Remark 1. The assumption that $\{v_{i,\tau}\}$ are i.i.d and independent of $\{\mathbf{x}_{it}\}$ seems strong in this setting. However, this is a common assumption for PDQ-CRE model, see Abrevaya and Dahl (2008), Harding and Lamarche (2017), and Cai, Chen, and Fang (2018), and for the quantile regression with measurement error, e.g., Hausman et al. (2021). The estimation method developed in Abrevaya and Dahl (2008), where the i.i.d. assumption is redundant, just ignored the influence of the unobserved disturbance $\{v_i\}$, which may cause the method to lose its validity, although it is simpler. To this end, we take the disturbance into account in this article. If the above independence assumption is not satisfied, it becomes very complicated. So, it is left as a future research topic.

Generally speaking, the integral in (4) does not have a close form unless it is assumed that $v_{i,\tau}$ has some particular density, such as the normality assumption as in Cai, Chen, and Fang (2018). Nevertheless, without the normality assumption of $v_{i,\tau}$, we propose using the Laplace approximation¹ method to approximate the above integral as in Breslow and Clayton (1993) and Harding and Hausman (2007). The Laplace approximation is a two-term Taylor expansion on the log density function. If \tilde{v} denotes the maxima of a density function $f(v)$, then it is also the maxima of the log probability density function $p(v) = \log f(v)$ and because $\dot{p}(\tilde{v}) = 0$, we can write:

$$\begin{aligned} p(v) &\simeq p(\tilde{v}) + (v - \tilde{v})\dot{p}(\tilde{v}) + \frac{1}{2}(v - \tilde{v})^2\ddot{p}(\tilde{v}) = p(\tilde{v}) + 0 + \frac{1}{2}(v - \tilde{v})^2\ddot{p}(\tilde{v}) \\ &= c + \frac{1}{2}(v - \tilde{v})^2\ddot{p}(\tilde{v}) = c - \frac{(v - a)^2}{2b^2}, \end{aligned}$$

¹For any probability density function that is smooth and unimodal, it can be approximated by a normal density function, which is the so-called Laplace approximation. Indeed, it is widely used in the literature; see, for instance, the papers by Breslow and Clayton (1993) and Harding and Hausman (2007).

where $a = \tilde{v}$, $b = \{-\ddot{p}(\tilde{v})\}^{-1/2}$ with $\ddot{p}(\tilde{v}) < 0$ because \tilde{v} is a maxima, and $c = p(\tilde{v})$. By assuming the mean of the disturbance v_i is zero; see Assumption A2 in the Appendix, thus $f(v)$ can be approximated by:

$$\tilde{f}(v) = c \cdot e^{-\frac{v^2}{2\sigma^2}},$$

where σ^2 is the variance of v_i . Since $\tilde{f}(v)$ is actually a probability density function, c is a function of σ with $c = 1/\sqrt{2\pi}\sigma$. Therefore, the approximation yields

$$l(y_{it}, q_{\tau, -v}(y_{it}|\mathbf{x}_i)) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} l(y_{it}, q_{\tau}(y_{it}|\mathbf{x}_i, v_{i,\tau})) e^{-\frac{v_{i,\tau}^2}{2\sigma^2}} dv_{i,\tau}. \tag{5}$$

According to Komunjer (2005), different choices of $l(y, q)$ would affect the asymptotic theory of QMLE for quantile models. In this article, for simplicity, we define $l(y, q)$ as

$$l(y, q) = e^{-\rho_{\tau}(y-q)}, \tag{6}$$

which is a probability density function, which is a member of the so-called tick-exponential family; see Komunjer (2005) for details. Substituting (6) into the integrated quasi-likelihood function for the τ th conditional quantile in (5), we have

$$l(y_{it}, q_{\tau, -v}(y_{it}|\mathbf{x}_i)) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\rho_{\tau}(y_{it} - q_{\tau}(y_{it}|\mathbf{x}_i, v_{i,\tau})) - \frac{v_{i,\tau}^2}{2\sigma^2}\right\} dv_{i,\tau}.$$

By a simple calculation, we get

$$l_{\tau}(u_{it}, \sigma) \equiv l(y_{it}, q_{\tau, -v}(y_{it}|\mathbf{x}_i)) = \exp\{-\rho_{\tau}(u_{it})\} v_{\tau}(u_{it}, \sigma) (I_{u_{it} \geq 0} + \exp\{-u_{it}\} I_{u_{it} < 0}), \tag{7}$$

where $u_{it} = y_{it} - q_{\tau}(y_{it}|\mathbf{x}_i, v_{i,\tau})$, $v_{\tau}(u_{it}, \sigma) = \exp\{\tau^2\sigma^2/2\}\Phi(u_{it}/\sigma - \tau\sigma) + \exp\{(\tau - 1)^2\sigma^2/2\}\Phi(-u_{it}/\sigma + (\tau - 1)\sigma) \exp\{u_{it}\}$, and $\Phi(\cdot)$ is the standard normal distribution function. Therefore, the log quasi-likelihood function $L_{\tau}(\xi_{\tau}, \sigma)$ is given by

$$\begin{aligned} L_{\tau}(\xi_{\tau}, \sigma) &= \sum_{i=1}^N \sum_{t=1}^T \ln l_{\tau}(u_{it}, \sigma) \\ &= \sum_{i=1}^N \sum_{t=1}^T [-\rho_{\tau}(u_{it}) + \ln(v_{\tau}(u_{it}, \sigma)) + \ln(I_{u_{it} \geq 0} + \exp\{-u_{it}\} I_{u_{it} < 0})] \\ &= \sum_{i=1}^N \sum_{t=1}^T [-\rho_{\tau}(u_{it}) + \ln(v_{\tau}(u_{it}, \sigma)) - u_{it} I_{u_{it} < 0}] \\ &= \sum_{i=1}^N \sum_{t=1}^T [-\tau u_{it} + \ln(v_{\tau}(u_{it}, \sigma))]. \end{aligned}$$

Hence, for a given $\tilde{\sigma}$, the Integrated Quasi-Maximum Likelihood Estimate (IQMLE) of ξ_{τ} is obtained by

$$\hat{\xi}_{\tau} = \arg \max_{\xi_{\tau}} L_{\tau}(\xi_{\tau}, \tilde{\sigma}), \tag{8}$$

and the IQMLE of δ_{τ} is given by

$$\hat{\delta}_{\tau} = \mathbf{e}_1^T \hat{\xi}_{\tau}, \tag{9}$$

where \mathbf{e}_1 is a $(KT + K + 1) \times K$ matrix with the elements at (i, i) , where $1 \leq i \leq K$, are 1's and other elements are 0's. Clearly, the purpose of \mathbf{e}_1 is selecting the first K elements in the $(KT + K + 1)$ -dimension $\hat{\xi}_{\tau}$. In practical implementation, we can estimate both ξ_{τ} and σ by iterating $\hat{\xi}_{\tau} = \arg \max_{\xi_{\tau}} L_{\tau}(\xi_{\tau}, \tilde{\sigma})$ and $\tilde{\sigma} = \arg \max_{\sigma} L_{\tau}(\hat{\xi}_{\tau}, \sigma)$ until convergence.

Remark 2. Actually, if we set $\tilde{\sigma} = 0$ in solving (8) and (9), we just obtain the same estimation as Abrevaya and Dahl (2008), since the log quasi-likelihood function now degenerates to the form $L_\tau(\xi_\tau, 0) = \sum_{i=1}^N \sum_{t=1}^T -\rho_\tau(u_{it})$. To this end, our proposed IQMLE contains the case of the classical Abrevaya and Dahl (2008) estimator.

2.3. Large sample theory

This section provides asymptotic theory of our proposed estimator given in Section 2.2. All assumptions and proofs are relegated to the Appendix. Let $\hat{\delta}_\tau$ be the IQMLE of δ_τ obtained by solving (8) and (9). Here follows the weak consistency of our proposed estimator.

Theorem 1. (Consistency). *Suppose that Assumptions A1–A4 in the Appendix and $T \geq 2$ hold. Then, we have $\hat{\delta}_\tau \xrightarrow{P} \delta_\tau$ as $N \rightarrow \infty$ with \sqrt{N} -rate, where \xrightarrow{P} denotes the convergence in probability.*

Some additional notations are defined here for the following asymptotic theorem. To this end, we let $\mathbf{g}_{it} = (1, \mathbf{x}_{it}^\top, \mathbf{x}_{i1}^\top, \dots, \mathbf{x}_{iT}^\top)^\top$, $u = y - q_\tau(y|\mathbf{x})$, $\psi_\tau(u, \sigma) = -\partial \ln l_\tau(u, \sigma) / \partial u \equiv \tau - h_\tau(u, \sigma)$, where $h_\tau(u, \sigma) = \partial \ln(v_\tau(u, \sigma)) / \partial u$, and $\dot{h}_\tau(u, \sigma) = \partial h_\tau(u, \sigma) / \partial u$. Also, define

$$\begin{aligned} D_{gh}(\sigma) &= \Sigma_0 E[-\dot{h}_\tau(y_{it} - q_\tau(y_{it}|\mathbf{x}_i), \sigma)], \\ D_g(\sigma) &= \Sigma_0 E[\psi_\tau(y_{it} - q_\tau(y_{it}|\mathbf{x}_i), \sigma)], \\ D_{g1t}(\sigma) &= \Sigma_{t-1} E[\psi_\tau(y_{i1} - q_\tau(y_{i1}|\mathbf{x}_i), \sigma) \psi_\tau(y_{it} - q_\tau(y_{it}|\mathbf{x}_i), \sigma)], \end{aligned}$$

where $\Sigma_t = E(\mathbf{g}_{i1} \mathbf{g}_{it+1}^\top)$ for $t \geq 0$.

Theorem 2. (Asymptotic Normality). *Under the assumptions of Theorem 1, we have*

$$\sqrt{N}(\hat{\delta}_\tau - \delta_\tau) \xrightarrow{d} \mathcal{N}(0, \Sigma(\sigma, T)), \quad \text{as } N \rightarrow \infty,$$

where $\Sigma(\sigma, T) = \mathbf{e}_1^\top D_{gh}^{-1}(\sigma) \{D_g(\sigma) / T + \sum_{t=2}^T 2(1/T - t/T^2 + 1/T^2) D_{g1t}(\sigma)\} D_{gh}^{-1}(\sigma) \mathbf{e}_1$, and \xrightarrow{d} denotes the convergence in distribution.

Remark 3. The asymptotic distribution theory in Theorem 2 permits T is fixed and covers the case of non zero dependence over time through the term $D_{g1t}(\sigma)$. By Assumption A4 and Proposition 1 in the Appendix, $D_{g1t}(\sigma)$ can be bounded for any given σ . Given more assumptions on the temporal dependence of $\{y_{it}, \mathbf{x}_{it}\}$, one can obtain the limit form of $\Sigma(\sigma, T)$ when $T \rightarrow \infty$. While, this is not the main focus of this article.

We can observe that the covariance-variance matrix $\Sigma(\sigma, T)$ is a function of both T and σ . When $\sigma = 0$, the model degenerates to the case that $v_{i,\tau} \equiv 0$. To be specific, when $v_{i,\tau} \equiv 0$, i.e., $\sigma = 0$,

$$h_\tau(u, 0) = \lim_{\sigma \rightarrow 0^+} \frac{\partial \ln(v_\tau(u, \sigma))}{\partial u} = \lim_{\sigma \rightarrow 0^+} \frac{e^{u + \frac{1-2\tau}{2}\sigma^2} \Phi(-\frac{u}{\sigma} + (\tau - 1)\sigma)}{e^{u + \frac{1-2\tau}{2}\sigma^2} \Phi(-\frac{u}{\sigma} + (\tau - 1)\sigma) + \Phi(\frac{u}{\sigma} - \tau\sigma)} = I_{u < 0},$$

and

$$D_{gh}(0) = \Sigma_0 E[-\dot{h}_\tau(y_{it} - q_\tau(y_{it}|\mathbf{x}_i), 0)] = \Sigma_0 f_y(F^{-1}(\tau)),$$

where $f_y(\cdot)$ is the density function of dependent variable y_{it} and $F(\cdot)$ is its CDF. The second equality follows from (A.7)-(A.9) in the Appendix if we use the Knight identity in Knight (1998) for cross-sectional data; see the detailed proof in the Appendix. Then, we can rewrite the asymptotic covariance

in [Theorem 2](#) for this special case as follows

$$\begin{aligned} \Sigma(0, T) &= \mathbf{e}_1^T f_y^{-2}(F^{-1}(\tau)) \Sigma_0^{-1} \left\{ \tau(1 - \tau) \Sigma_0 \right. \\ &\quad \left. + \sum_{t=2}^T 2(1/T - t/T^2 + 1/T^2) \Sigma_{t-1} E[\psi_\tau(y_{i1} - q_\tau(y_{i1} | \mathbf{x}_i), \sigma) \psi_\tau(y_{it} - q_\tau(y_{it} | \mathbf{x}_i), \sigma)] \right\} \Sigma_0^{-1} \mathbf{e}_1, \end{aligned}$$

Clearly, $\Sigma(0, T)$ is the asymptotic covariance with no disturbance $\{v_i\}$. The existence of the second term $\sum_{t=2}^T 2(1/T - t/T^2 + 1/T^2) \Sigma_{t-1} E[\psi_\tau(u_{i1}, \sigma) \psi_\tau(u_{it}, \sigma)]$ is because we allow $\{\mathbf{x}_{it}\}$ to be correlated over t ; see [Assumption A1](#) in the Appendix. When $T = 1$,

$$\Sigma(0, 1) = \mathbf{e}_1^T \tau(1 - \tau) f_y^{-2}(F^{-1}(\tau)) \Sigma_0^{-1} \Sigma_0 \Sigma_0^{-1} \mathbf{e}_1 = \mathbf{e}_1^T \tau(1 - \tau) f_y^{-2}(F^{-1}(\tau)) \Sigma_0^{-1} \mathbf{e}_1,$$

which is the asymptotic covariance matrix for classical linear quantile model for i.i.d. cross sectional data. Therefore, model (3) and its corresponding estimator can be viewed as an extension of quantile regression from cross sectional data to panel data.

2.4. Hypothesis testing

To make a statistical test, one needs a consistent estimate of the asymptotic covariance matrix $\Sigma(\sigma, T)$. The explicit expression of $\Sigma(\sigma, T)$ in [Theorem 2](#) provides a natural consistent estimator. We will show that $\hat{\Sigma}(\sigma, T)$ is a consistent estimator of $\Sigma(\sigma, T)$ for any given σ . First, define

$$\begin{aligned} \hat{D}_{gh}(\sigma) &= (NT)^{-2} \sum_{i=1}^N \sum_{t=1}^T \mathbf{X}_{0,it} \cdot \sum_{i=1}^N \sum_{t=1}^T -\dot{h}_\tau(y_{it} - \hat{y}_{it,\tau}, \sigma), \\ \hat{D}_g(\sigma) &= (NT)^{-2} \sum_{i=1}^N \sum_{t=1}^T \mathbf{X}_{0,it} \cdot \sum_{i=1}^N \sum_{t=1}^T \psi_\tau^2(y_{it} - \hat{y}_{it}, \sigma), \\ \hat{D}_{g1t}(\sigma) &= (N(T - t + 1))^{-2} \sum_{i=1}^N \sum_{s=1}^{T-t+1} \mathbf{X}_{t,is} \cdot \sum_{i=1}^N \sum_{s=1}^{T-t+1} \psi_\tau(y_{is} - \hat{y}_{is,\tau}, \sigma) \psi_\tau(y_{i,s+t} - \hat{y}_{i,s+t,\tau}, \sigma), \end{aligned}$$

where $\hat{y}_{it,\tau} = \hat{\xi}_\tau^T \mathbf{g}_{it}$, $\mathbf{X}_{0,it} = \mathbf{g}_{it} \mathbf{g}_{it}^T$, and $\mathbf{X}_{t,is} = \mathbf{g}_{is} \mathbf{g}_{i,s+t}^T$. It can be easily shown that $\hat{D}_g(\sigma) = D_g(\sigma) + o_p(1)$, $\hat{D}_{g1t}(\sigma) = D_{g1t}(\sigma) + o_p(1)$, and $\hat{D}_{gh}(\sigma) = D_{gh}(\sigma) + o_p(1)$ by the law of large numbers. Then, the consistent estimator of $\Sigma(\sigma, T)$ is given by

$$\hat{\Sigma}(\sigma, T) = \mathbf{e}_1^T \hat{D}_{gh}^{-1}(\sigma) \left\{ \hat{D}_g(\sigma) + \sum_{t=2}^T 2(1/T - t/T^2 + 1/T^2) \hat{D}_{g1t}(\sigma) \right\} \hat{D}_{gh}^{-1}(\sigma) \mathbf{e}_1. \quad (10)$$

However, we still do not know the value of nuisance parameter σ . We will replace it by $\tilde{\sigma}$, which is obtained by iterating $\tilde{\xi}_\tau = \arg \max_{\xi_\tau} L_\tau(\xi_\tau, \tilde{\sigma})$ and $\tilde{\sigma} = \arg \max_{\sigma} L_\tau(\tilde{\xi}_\tau, \sigma)$ until convergence for statistical inference.

To test $H_0 : \mathbf{R}\delta_\tau = r_\tau$ (including the simple test $H_0 : \delta_{j,\tau} = 0$ for each j , which is implemented in [Section 4](#) in our empirical study for computing p -value), where \mathbf{R} is a $J \times K$ known matrix with the rank J and r_τ is a known constant, a Wald type test statistic can be constructed as follows:

$$W_N = N(\mathbf{R}\hat{\delta}_\tau - r_\tau)^T \left[\mathbf{R}\hat{\Sigma}(\tilde{\sigma}, T)\mathbf{R}^T \right]^{-1} (\mathbf{R}\hat{\delta}_\tau - r_\tau).$$

The limiting distribution of W_N is stated in the following theorem, which can be established easily by [Theorem 2](#) and Slutsky's theorem and omitted.

Theorem 3. (*Asymptotic χ^2 Test*). Under the assumptions of [Theorem 1](#) and the null hypothesis $H_0 : \mathbf{R}\delta_\tau = r_\tau$, we have

$$W_N \xrightarrow{d} \chi_J^2, \quad \text{as } N \rightarrow \infty,$$

where χ_J^2 is the χ^2 -distribution with J degrees of freedom.

3. A Monte Carlo simulation study

In this section, we examine the finite sample performance of the proposed estimator of our model. We consider the following data generating process: for $1 \leq i \leq N, 1 \leq t \leq N$,

$$y_{it} = \delta_0 + \delta_1 x_{it,1} + \delta_2 x_{it,2} + \alpha_i + (0.5 + 0.3x_{it,1} + 0.1x_{it,2})e_{it}, \tag{11}$$

where $\alpha_i = \sum_{t=1}^T \eta_{t,1}x_{it,1} + \eta_{t,2}x_{it,2} + v_i$ and T is relatively small compared to the sample size N . Here, the sequence $\{e_{it}\}$ is generated from $\mathcal{N}(0, 1)$ and the covariates $x_{it,1}$ and $x_{it,2}$ are both generated from i.i.d. $U(0, 2)$. Furthermore, the measurement error (disturbance) $\{v_i\}$ is a sequence of i.i.d. random variables generated from $\mathcal{N}(0, 0.1^2)$, $\mathcal{N}(0, 0.3^2)$, and Laplace(0, 0.2) distribution to check the robustness. The parameters are set by $\delta_0 = 3, \delta_1 = -1, \delta_2 = 1$, and η_{ij} are generated from i.i.d. $U(-0.3, 0.3)$. Therefore, the quantile model of the DGP in (11) is

$$q_\tau(y_{it} | \mathbf{x}_i, v_i) = \delta_{0,\tau} + \delta_{1,\tau}x_{it,1} + \delta_{2,\tau}x_{it,2} + \alpha_i,$$

where $\delta_{0,\tau} = 3 + 0.5\Phi^{-1}(\tau), \delta_{1,\tau} = -1 + 0.3\Phi^{-1}(\tau), \delta_{2,\tau} = 1 + 0.1\Phi^{-1}(\tau)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, and $\alpha_i = \sum_{t=1}^T \eta_{t,1}x_{it,1} + \eta_{t,2}x_{it,2} + v_i$. The individual effect $\alpha_{i,\tau}$ in model (1) is actually $\delta_{0,\tau} + \alpha_i$ here. Following, we show the finite sample performance of our proposed method in both parameter estimation and statistical inference.

3.1. Estimation

To evaluate the performance of $\hat{\delta}_{j,\tau}$, for $0 \leq j \leq 2$, we use the root mean square error (RMSE) as the criterion, which is defined as

$$\text{RMSE}(\hat{\delta}_{j,\tau}) = \sqrt{\frac{1}{B} \sum_{b=1}^B \left(\hat{\delta}_{j,\tau}^{(b)} - \delta_{j,\tau} \right)^2},$$

where B is the number of repetitions, and $\hat{\delta}_{j,\tau}^{(b)}$ is the estimation of $\delta_{j,\tau}$ in the b -th repetition. For comparison, we provide the RMSE's of three estimators: (i) Abrevaya and Dahl estimator proposed by Abrevaya and Dahl (2008), which is just the case $\tilde{\sigma} = 0$; (ii) IQMLE obtained by (8) and (9), where $\tilde{\sigma}$ is estimated by one-step iteration; (iii) IQMLE with priori knowledge, where $\tilde{\sigma}$ is set as its true value. We denote these three estimators as ADE, IQMLE, and IQMLEw, respectively. We consider the sample size $N = 40, 60, 80, \dots, 300$, the time period $T = 2$ and 10, and the quantile $\tau = 0.2, 0.5$, and 0.8. For each setting, we conduct the simulation $B = 1000$ times. The RMSE's for each setting are shown in Figures 1 ($T = 2$ and $\{v_i\}$ is generated form $\mathcal{N}(0, 0.3^2)$), 2 ($T = 2$ and $\{v_i\}$ is generated form Laplace(0, 0.2)), 3 ($T = 10$ and $\{v_i\}$ is generated form $\mathcal{N}(0, 0.3^2)$), and 4 $T = 10$ and $\{v_i\}$ is generated form Laplace(0, 0.2)), respectively.

As seen from these figures, on the one hand, the RMSE's of the three estimators significantly decrease when the sample size N increases under different lengths of period T and generating distributions of $\{v_i\}$. For example, in the case of $T = 2, v_i \sim \mathcal{N}(0, 0.3^2)$ and $\tau = 0.5$, when the sample size increases from 40 to 300, the RMSEs of IQMLE for $\hat{\delta}_{0,\tau}, \hat{\delta}_{1,\tau}$ and $\hat{\delta}_{2,\tau}$ decrease quickly from 0.4193 to 0.1413, from 0.2772 to 0.0945, and from 0.2574 to 0.0908, respectively. In the case of $T = 10, v_i \sim \text{Laplace}(0, 0.2)$ and $\tau = 0.2$, the RMSEs of IQMLE for $\hat{\delta}_{0,\tau}, \hat{\delta}_{1,\tau}$ and $\hat{\delta}_{2,\tau}$ decrease quickly from 0.8234 to 0.2381, from 0.1095 to 0.0521, and from 0.1022 to 0.0373, respectively.

On the other hand, our proposed estimators IQMLE and IQMLEw are superior to ADE in most cases. Actually, ADE is a special case of IQMLE, which presumes that the variance of disturbance is zero. We also notice that in the cases for $T = 2, \tau = 0.2$, and 0.8, the convergence rate of IQMLE is relatively slower than the other two, which is due to the inaccurate estimation of σ since with priori knowledge, the performance of IQMLEw is pretty well. When T increases to 10, the estimation errors of IQMLE and IQMLEw are very close, thus the estimation problem of σ will be solved when T is large. In summary, the

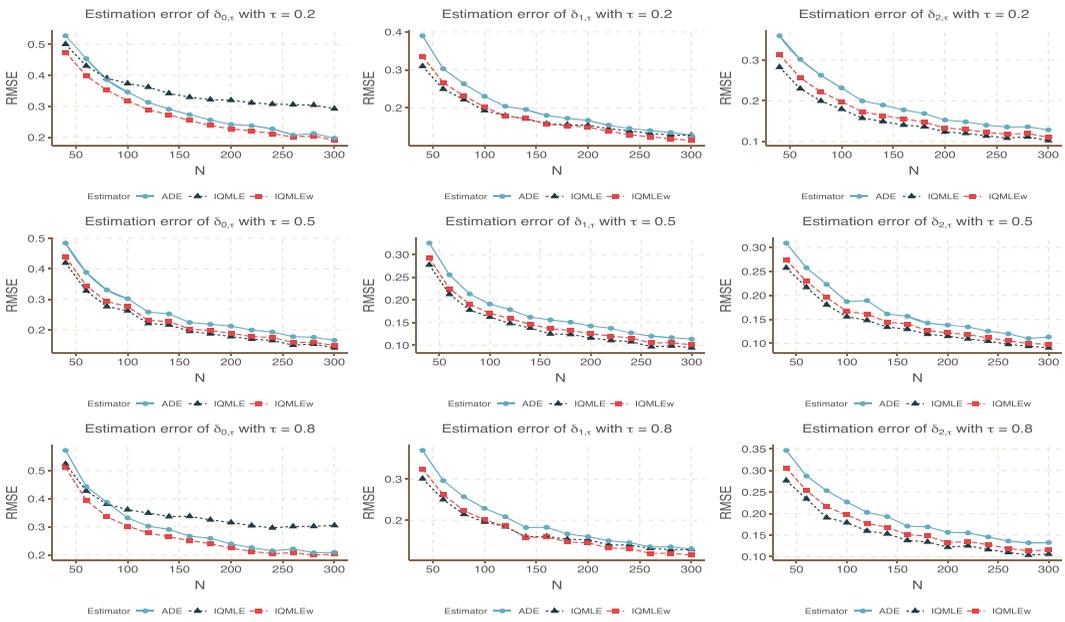


Figure 1. The RMSEs when $T = 2$ and $\{v_j\}$ is generated from $\mathcal{N}(0, 0.3^2)$.

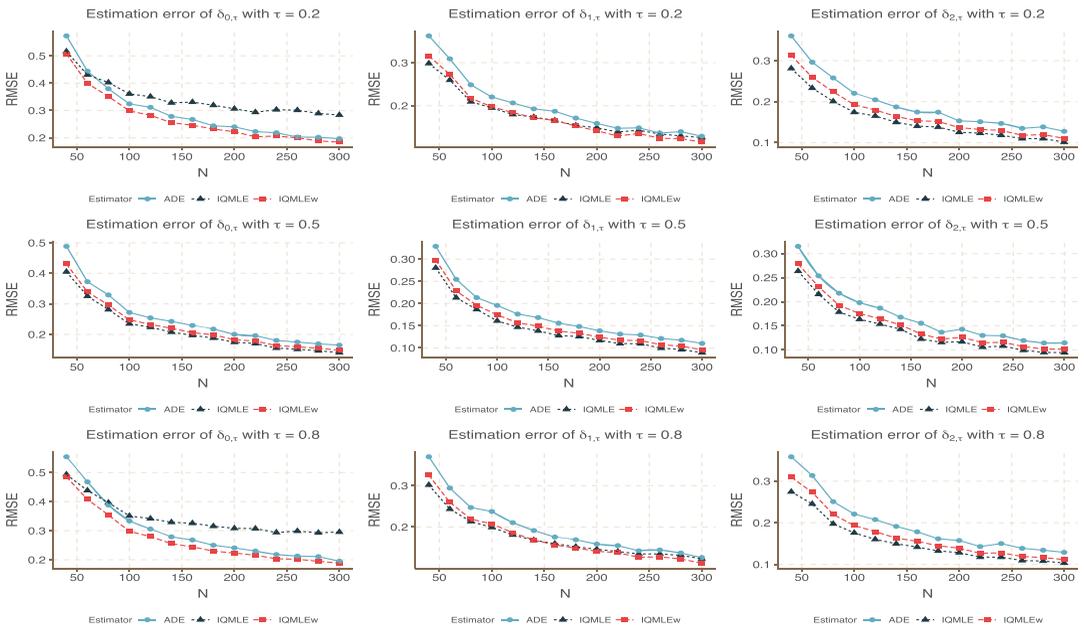


Figure 2. The RMSEs when $T = 2$ and $\{v_j\}$ is generated from Laplace(0, 0.2).

decrease of RMSEs with larger sample size and the superior performance compared with the Abrevaya and Dahl (2008) estimator indicate that our simulation results are consistent with the asymptotic theory established in Section 2.3.

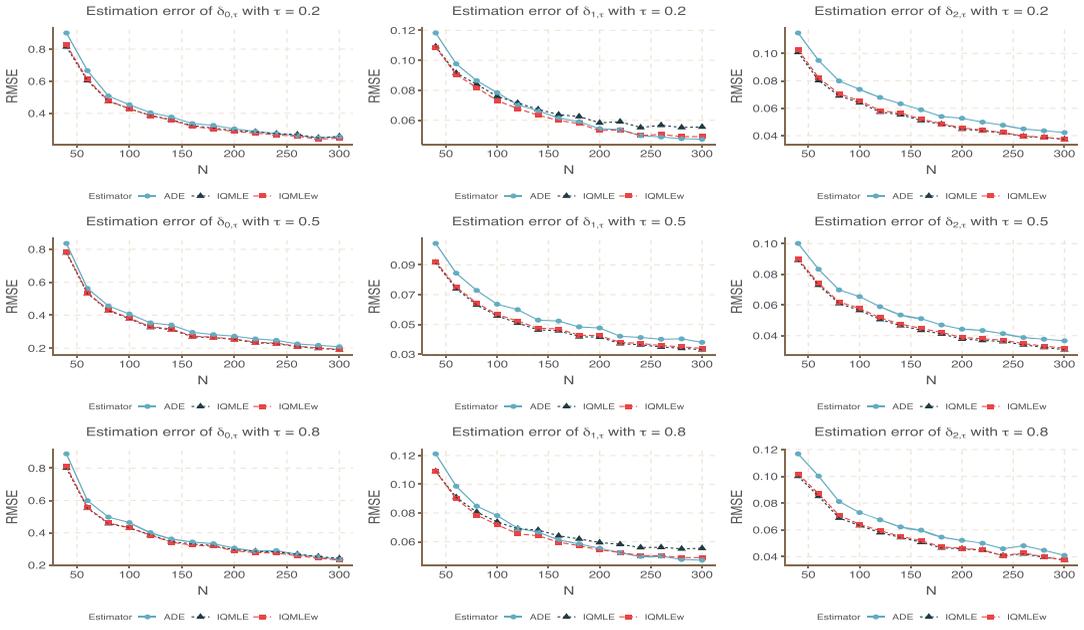


Figure 3. The RMSEs when $T = 10$ and $\{v_j\}$ is generated from $\mathcal{N}(0, 0.3^2)$.

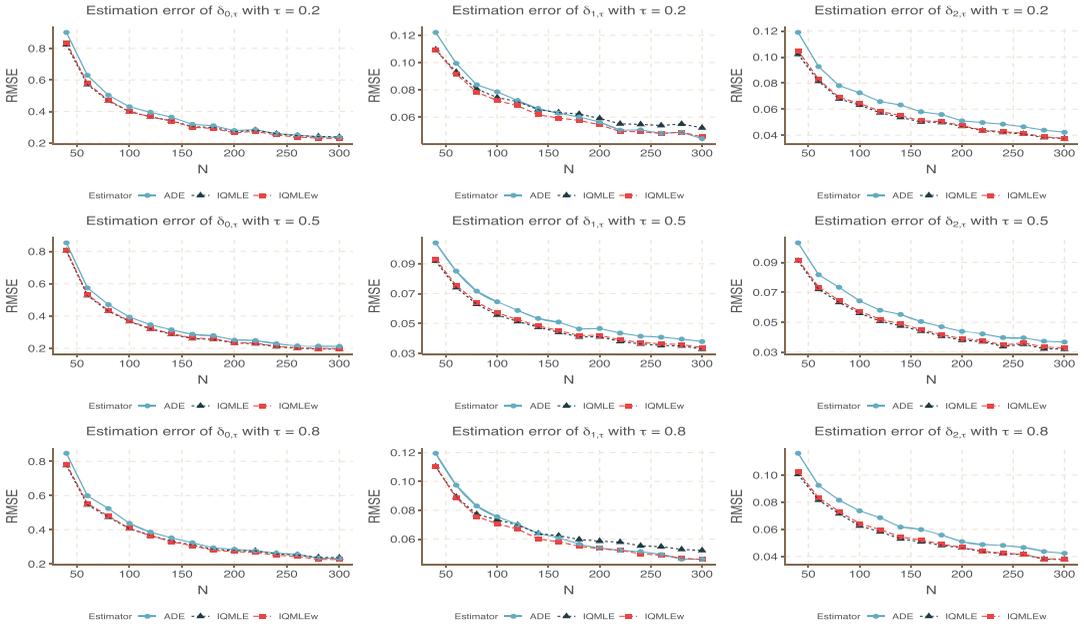


Figure 4. The RMSEs when $T = 10$ and $\{v_j\}$ is generated from Laplace(0, 0.2).

3.2. Inference

To illustrate the empirical significance level and statistical power of the test proposed in Section 2.4, we again use the DGP in (11). We consider the tests $H_0 : \mathbf{R}\delta_\tau = r_\tau$ v.s. $H_1 : \mathbf{R}\delta_\tau \neq r_\tau$, where the coefficient matrices \mathbf{R} are chosen as (0.3, 0.3, 0.4), (0.9, 0.05, 0.05), and (0.5, -1, 0.5) with rank $J = 1$ to check the robustness. Denote the true value of r_τ as $r_{\tau,0} \equiv \mathbf{R}\delta_\tau$ in each case. To obtain the empirical power, we consider the null hypothesis with $r_\tau = 1.1r_{\tau,0}$. The disturbance $\{v_j\}$ is generated from $\mathcal{N}(0, 0.1^2)$ and

Table 1. The empirical sizes and powers of the proposed tests.

N	τ	$R = (0.3, 0.3, 0.4)$			$R = (0.9, 0.05, 0.05)$			$R = (0.5, -1, 0.5)$		
		0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
Panel A: $v_i \sim \mathcal{N}(0, 0.1^2)$, $r_\tau = r_{\tau,0}$										
100		0.1238	0.1041	0.1220	0.1216	0.0988	0.1248	0.1099	0.0837	0.1141
200		0.0848	0.0737	0.0825	0.0776	0.0669	0.0763	0.0767	0.0634	0.0768
300		0.0704	0.0627	0.0709	0.0678	0.0599	0.0704	0.0692	0.0549	0.0631
400		0.0675	0.0612	0.0697	0.0620	0.0556	0.0612	0.0561	0.0520	0.0612
500		0.0627	0.0580	0.0618	0.0541	0.0507	0.0552	0.0574	0.0524	0.0550
600		0.0600	0.0579	0.0601	0.0566	0.0490	0.0573	0.0559	0.0503	0.0560
700		0.0598	0.0605	0.0611	0.0563	0.0510	0.0561	0.0586	0.0486	0.0541
800		0.0594	0.0571	0.0585	0.0546	0.0524	0.0558	0.0571	0.0499	0.0538
Panel B: $v_i \sim \text{Laplace}(0, 0.2)$, $r_\tau = r_{\tau,0}$										
100		0.0802	0.0736	0.0772	0.0764	0.0723	0.0850	0.0731	0.0601	0.0780
200		0.0684	0.0686	0.0657	0.0689	0.0568	0.0683	0.0686	0.0539	0.0713
300		0.0650	0.0605	0.0625	0.0683	0.0529	0.0644	0.0694	0.0476	0.0727
400		0.0605	0.0539	0.0608	0.0641	0.0466	0.0663	0.0796	0.0491	0.0792
500		0.0594	0.0566	0.0610	0.0683	0.0461	0.0669	0.0801	0.0464	0.0832
600		0.0609	0.0550	0.0589	0.0766	0.0485	0.0689	0.0880	0.0462	0.0807
700		0.0602	0.0504	0.0593	0.0741	0.0451	0.0749	0.0944	0.0436	0.0904
800		0.0600	0.0522	0.0568	0.0765	0.0481	0.0760	0.0972	0.0467	0.0936
Panel C: $v_i \sim \mathcal{N}(0, 0.1^2)$, $r_\tau = 1.1r_{\tau,0}$										
100		0.2111	0.2805	0.3395	0.2493	0.2964	0.3129	0.5061	0.5189	0.4487
200		0.2404	0.3879	0.4634	0.3114	0.4170	0.4114	0.7087	0.7591	0.6544
300		0.2990	0.5149	0.5889	0.3870	0.5362	0.5419	0.8540	0.8941	0.7953
400		0.3616	0.6149	0.6925	0.4633	0.6479	0.6433	0.9259	0.9646	0.8904
500		0.4175	0.7124	0.7705	0.5522	0.7468	0.7311	0.9641	0.9847	0.9485
600		0.4805	0.7805	0.8405	0.6248	0.8096	0.8067	0.9871	0.9943	0.9729
700		0.5360	0.8380	0.8888	0.6884	0.8659	0.8566	0.9949	0.9987	0.9870
800		0.5938	0.8814	0.9234	0.7365	0.9022	0.8972	0.9971	0.9996	0.9944
Panel D: $v_i \sim \text{Laplace}(0, 0.2)$, $r_\tau = 1.1r_{\tau,0}$										
100		0.1796	0.2418	0.2773	0.2272	0.2503	0.2265	0.5314	0.5014	0.3430
200		0.2454	0.3840	0.4275	0.3610	0.4073	0.3408	0.7927	0.7742	0.5671
300		0.3363	0.5180	0.5611	0.4786	0.5493	0.4707	0.9253	0.9128	0.7363
400		0.4177	0.6376	0.6799	0.5860	0.6619	0.5668	0.9766	0.9693	0.8461
500		0.4924	0.7336	0.7702	0.6753	0.7535	0.6623	0.9922	0.9888	0.9161
600		0.5671	0.8021	0.8308	0.7549	0.8335	0.7441	0.9973	0.9968	0.9548
700		0.6273	0.8628	0.8902	0.8197	0.8866	0.7992	0.9993	0.9984	0.9779
800		0.6867	0.9014	0.9292	0.8678	0.9189	0.8520	0.9998	0.9997	0.9882

Laplace(0, 0.2), the time period is set as $T = 5$, and the theoretical significance level of the test is chosen as $\alpha = 5\%$. We consider the sample size $N = 100, 200, \dots, 800$, and the quantile $\tau = 0.2, 0.5$, and 0.8 . For each setting, we conduct the simulation $B = 10,000$ times. The empirical sizes and powers of the proposed tests are reported in Table 1.

On the one hand, as shown in Panels A and B of Table 1, the empirical significance levels (sizes) of the proposed asymptotic χ^2 test decrease close to 0.05, the theoretical significance level, when the sample size N increases. For example, in the case of $R = (0.5, -1, 0.5)$, $v_i \sim \mathcal{N}(0, 0.1^2)$, when the sample size increases from 100 to 800, the empirical sizes of the proposed test for $\tau = 0.2, 0.5$ and 0.8 decrease quickly from 0.1099 to 0.0571, from 0.0837 to 0.0499, and from 0.1141 to 0.0538, respectively.

On the other hand, as seen in Panels C and D, the empirical power increases as the sample size increases. For example, in the case of $R = (0.9, 0.05, 0.05)$, $v_i \sim \text{Laplace}(0, 0.2)$, when the sample size increases from 100 to 800, the empirical powers of the proposed test for $\tau = 0.2, 0.5$ and 0.8 increase quickly from 0.2272 to 0.8678, from 0.2503 to 0.9189, and from 0.2265 to 0.8520, respectively. The simulation results confirm the correctness and effectiveness of the asymptotic χ^2 test proposed in Section 2.4.

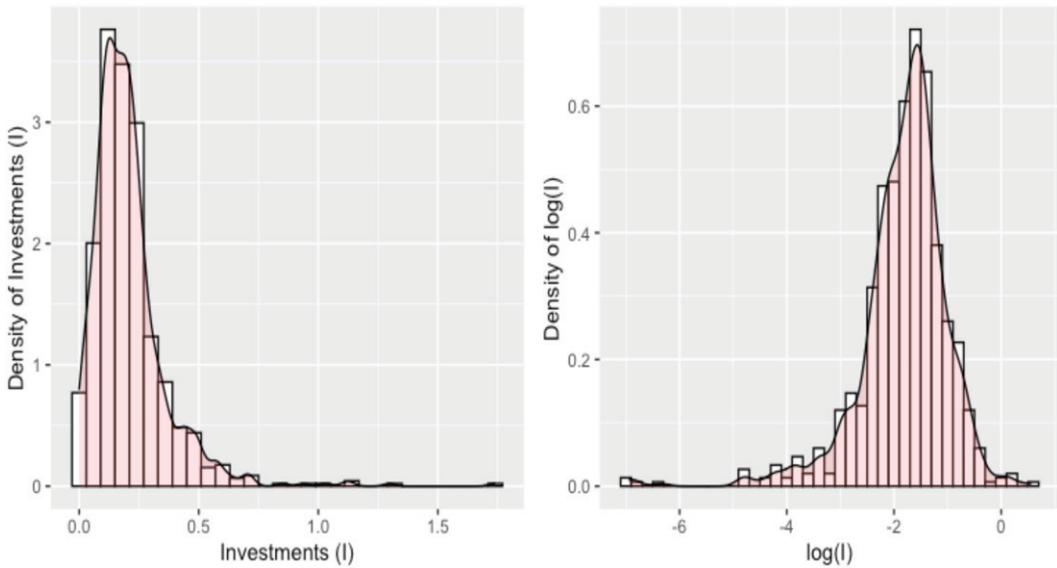


Figure 5. The estimated density of investment amount for year 2016: Left is for the original data and right is for the log-transformed data.

4. An empirical study for testing financial hierarchy

4.1. Empirical models

To illustrate our model empirically, we test the pecking order theory in the field of corporate finance, established by Myers (1984) and Myers and Majluf (1984). There are two important reasons for using the panel quantile regression to examine the pecking order theory. On the one hand, the prior literature employing mean models does not reach a consensus on whether corporate financing behaviors conform to the pecking order hypothesis (Brown et al., 2019). Some of prior studies confirm the pecking order theory analytically and empirically, indicating that firms prefer internal to external funds (i.e., the first rung of financial hierarchy), and prefer debt to equity issues (i.e., the second rung of financial hierarchy) when internal funds are exhausted and external funds are sought (Fama and French, 2002; Lemmon and Zender, 2010; Myers, 1984; Myers and Majluf, 1984; Shyam-Sunder and Myers, 1999). Nevertheless, the existing studies also document departures of financing hierarchy from the pecking order theory in different settings (Bharath, Pasquariello, and Wu, 2009; Brennan and Kraus, 1987; Brown, Fazzari, and Peterson, 2009; Fama and French, 2005; Frank and Goyal, 2003; Jung, Kim, and Stulz, 1996). Their either-or conclusions can be caused by the biased estimation of mean models. In contrast to the quantile regressions, their use of mean regressions does not take into consideration the fact that the distribution of the dependent variable is highly skewed. For instance, the distribution of corporate investments is highly skewed, which can be observed evidently from Figure 5 (the left is for the original data and the right for the log-transformed data) for the investment amount in year 2016. In addition, the employment of mean regressions fails to capture the phenomenon that the contribution of a given source of funds to financing investments could change across different levels of investments (Chay et al., 2015). On the other hand, Chay et al. (2015) used quantile regressions to test the pecking order hypothesis but they did not control firm and year fixed effects due to difficulties in theoretical modeling and computational challenges. Since it is difficult to identify valid instrumental variables for financing variables, the control of unobserved firm-specific heterogeneity, taken as a principal improvement of panel quantile regressions, is therefore critical for the examination of the pecking order theory.

Following Brown, Fazzari, and Peterson (2009) and Chay et al. (2015), we use a modified investment model which captures the responsiveness of investments to internal and external funds:

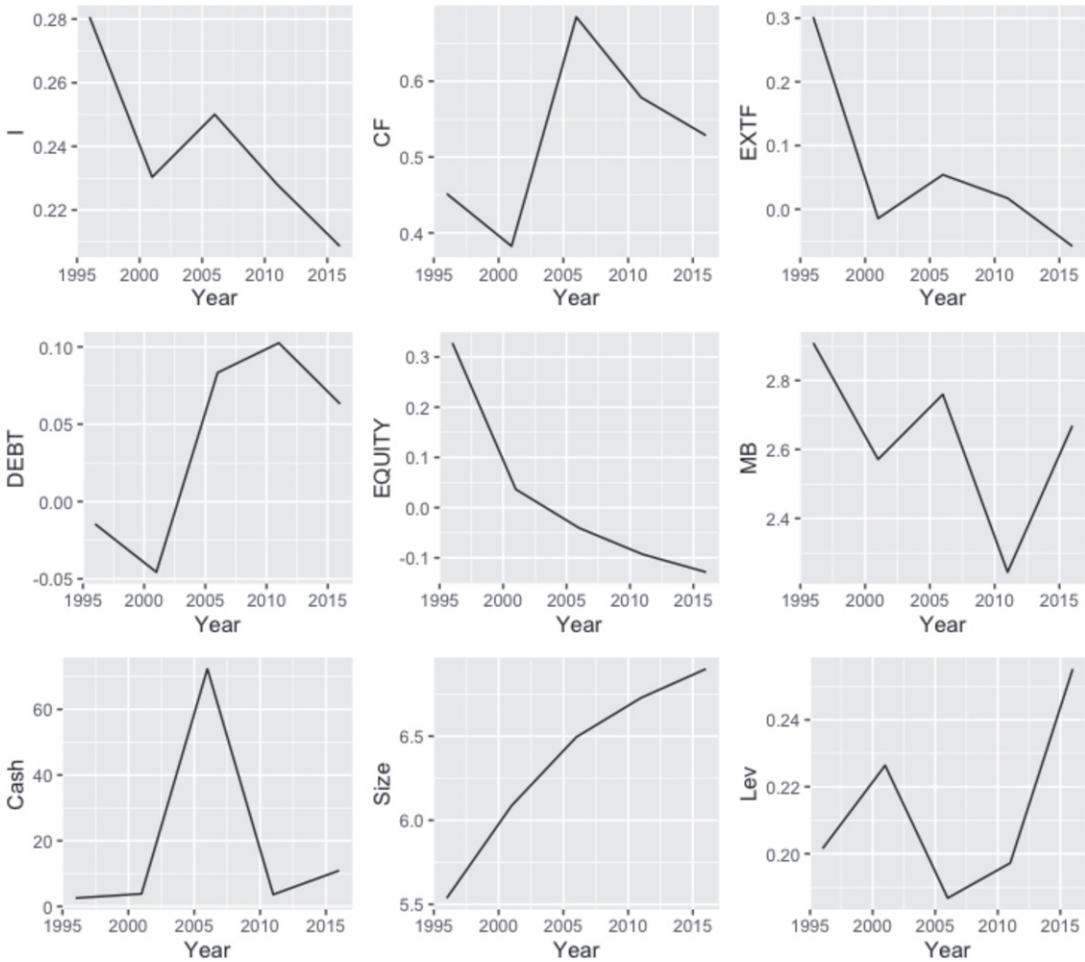


Figure 6. Time series of mean values for responsive and explanatory variables in each year.

$$q_{\tau}(I_{it}) = \alpha_{i,\tau} + \delta_{1,\tau}CF_{it} + \delta_{2,\tau}EXTF_{it} + \sum_{k=1}^K \gamma_{k,\tau}Control_{it,k}, \tag{12}$$

where I_{it} denotes the corporate investments in operating activities, measured as capital expenditures of firm i in fiscal year t . Concerning the key independent variables, internal funds (CF) is defined as the amount of cash flow and external funds (EXTF) is defined as the sum of debt issues (DEBT) and equity issues (EQUITY). All of the variables mentioned above are scaled by net fixed assets. Market-to-book value (MB), cash holdings (Cash), firm size (Size), leverage (Lev) and year dummy variables $\{d_t\}$ are also added as control variables to the regression above, and to all of the other regressions. Year dummy variables $\{d_t\}$ is set for $t = 2001, 2006, 2011,$ and 2016 since we set 1996 as the base year. We calculate the mean values of all responsive and explanatory variables in each year; see Figure 6. Indeed, we can observe there are significant differences in sample data across different years, which is consistent with the regression results of year dummies presented in Table 2. $\alpha_{i,\tau}$ is the individual effect used to control the unobserved firm-specific heterogeneity. In this article, we approximate $\alpha_{i,\tau}$, as discussed earlier, by a projection on the explanatory variables of firm i . Hence, the relative importance of internal and external fund can be observed from the magnitude of $\delta_{1,\tau}$ and $\delta_{2,\tau}$.

Table 2. Baseline tests of pecking order.

	Selected quantiles of investments (<i>I</i>)									
	OLS	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Panel A: First rung of pecking order										
CF	0.148 (0.018)***	0.086 (0.022)***	0.135 (0.021)***	0.156 (0.022)***	0.205 (0.023)***	0.224 (0.025)***	0.186 (0.027)***	0.19 (0.03)***	0.191 (0.036)***	0.271 (0.048)***
EXTF	0.068 (0.016)***	0.042 (0.016)***	0.026 (0.015)*	0.052 (0.016)***	0.041 (0.017)**	0.062 (0.018)***	0.071 (0.02)***	0.077 (0.021)***	0.069 (0.026)***	0.117 (0.035)***
MB	0.104 (0.016)***	0.054 (0.017)***	0.072 (0.016)***	0.09 (0.017)***	0.077 (0.018)***	0.064 (0.02)***	0.055 (0.021)**	0.059 (0.023)***	0.055 (0.028)**	0.1 (0.038)***
Cash	0.129 (0.02)***	-0.043 (0.025)*	0.027 (0.025)	0.063 (0.026)**	0.094 (0.027)***	0.116 (0.029)***	0.141 (0.032)***	0.176 (0.034)***	0.221 (0.042)***	0.192 (0.056)***
Size	-0.176 (0.053)***	0.136 (0.183)	0.074 (0.179)	-0.005 (0.189)	-0.068 (0.196)	-0.125 (0.211)	-0.203 (0.232)	-0.288 (0.251)	-0.385 (0.301)	-0.566 (0.403)
Lev	-0.117 (0.021)***	-0.1 (0.028)***	-0.124 (0.027)***	-0.083 (0.029)***	-0.1 (0.03)***	-0.081 (0.032)**	-0.096 (0.035)***	-0.125 (0.038)***	-0.134 (0.046)***	-0.075 (0.061)
d_{2001}	-0.184 (0.042)***	-0.111 (0.05)**	-0.155 (0.049)***	-0.16 (0.051)***	-0.162 (0.054)***	-0.169 (0.058)***	-0.19 (0.063)***	-0.181 (0.068)***	-0.205 (0.082)**	-0.199 (0.111)*
d_{2006}	-0.131 (0.046)***	-0.098 (0.059)*	-0.123 (0.057)**	-0.129 (0.061)**	-0.136 (0.064)**	-0.111 (0.069)	-0.113 (0.075)	-0.133 (0.081)	-0.141 (0.098)	-0.057 (0.132)
d_{2011}	-0.198 (0.049)***	-0.18 (0.067)***	-0.189 (0.065)***	-0.19 (0.069)***	-0.2 (0.072)***	-0.183 (0.078)**	-0.185 (0.084)**	-0.172 (0.091)*	-0.203 (0.11)*	-0.106 (0.149)
d_{2016}	-0.27 (0.051)***	-0.224 (0.073)***	-0.233 (0.071)***	-0.254 (0.075)***	-0.239 (0.078)***	-0.212 (0.085)**	-0.235 (0.092)**	-0.239 (0.099)**	-0.293 (0.12)**	-0.23 (0.162)
NT	3785	3785	3785	3785	3785	3785	3785	3785	3785	3785
Panel B: Second rung of pecking order										
CF	0.147 (0.019)***	0.104 (0.022)***	0.146 (0.021)***	0.160 (0.022)***	0.205 (0.023)***	0.222 (0.025)***	0.194 (0.028)***	0.185 (0.030)***	0.219 (0.036)***	0.27 (0.050)***
DEBT	0.049 (0.015)***	0.022 (0.013)*	0.029 (0.013)**	0.031 (0.013)**	0.029 (0.014)**	0.051 (0.016)***	0.058 (0.017)***	0.051 (0.018)***	0.075 (0.022)***	0.087 (0.031)***
EQUITY	0.044 (0.017)**	0.012 (0.019)	0.010 (0.018)	0.034 (0.019)*	0.028 (0.020)	0.055 (0.022)**	0.058 (0.024)**	0.075 (0.026)***	0.045 (0.031)	0.060 (0.044)
MB	0.104 (0.016)***	0.044 (0.017)***	0.087 (0.016)***	0.088 (0.017)***	0.077 (0.018)***	0.058 (0.020)***	0.061 (0.021)***	0.061 (0.023)***	0.063 (0.027)**	0.113 (0.039)***
Cash	0.130 (0.020)***	-0.016 (0.025)	0.025 (0.024)	0.091 (0.026)***	0.104 (0.027)***	0.124 (0.029)***	0.132 (0.032)***	0.192 (0.034)***	0.232 (0.041)***	0.246 (0.058)***
Size	-0.173 (0.054)***	0.142 (0.185)	0.074 (0.178)	-0.010 (0.189)	-0.071 (0.198)	-0.132 (0.215)	-0.209 (0.236)	-0.254 (0.252)	-0.412 (0.301)	-0.491 (0.422)
Lev	-0.117 (0.021)***	-0.089 (0.028)***	-0.129 (0.027)***	-0.078 (0.028)***	-0.096 (0.030)***	-0.091 (0.033)***	-0.099 (0.036)***	-0.131 (0.038)***	-0.143 (0.046)***	-0.073 (0.064)
d_{2001}	-0.186 (0.042)***	-0.125 (0.049)**	-0.154 (0.048)***	-0.161 (0.050)***	-0.172 (0.053)***	-0.175 (0.058)***	-0.202 (0.063)***	-0.195 (0.067)***	-0.157 (0.081)*	-0.157 (0.114)
d_{2006}	-0.133 (0.046)***	-0.115 (0.059)*	-0.129 (0.057)**	-0.127 (0.060)**	-0.133 (0.063)**	-0.124 (0.069)*	-0.108 (0.075)	-0.134 (0.079)*	-0.117 (0.096)	-0.074 (0.135)
d_{2011}	-0.202 (0.049)***	-0.203 (0.066)***	-0.194 (0.064)***	-0.196 (0.067)***	-0.206 (0.072)***	-0.186 (0.078)**	-0.182 (0.085)**	-0.191 (0.090)**	-0.148 (0.108)	-0.135 (0.153)
d_{2016}	-0.273 (0.051)***	-0.244 (0.072)***	-0.239 (0.070)***	-0.271 (0.073)***	-0.248 (0.078)***	-0.206 (0.084)**	-0.233 (0.092)**	-0.260 (0.097)***	-0.238 (0.118)**	-0.273 (0.166)***
NT	3785	3785	3785	3785	3785	3785	3785	3785	3785	3785

Table 2 reports the results of tests on pecking order theory employing the OLS regression and the panel quantile regression. Panel A concerning on the first rung of the pecking order reports the estimates of the internal and external funds on investment activities. Panel B concerning on the second order of pecking order reports the estimates of the internal funds, debt issues and equity issues, respectively. The dependent variable denotes the corporate investments in operating activities, measured as capital expenditures scaled by net fixed assets. The independent variable CF is defined as cash flow, scaled by net fixed assets. Another variable of interest is EXTF, measured as the sum of debt issues (DEBT) and equity issues (EQUITY), scaled by net fixed assets, where DEBT is calculated as funds raised by debt issues and EQUITY is calculated as funds raised by equity issues. We control for market-to-book value (MB), Cash holdings (Cash), firm size (Size), and leverage (Lev). Panel B reports the results of the internal funds, debt issues and equity issues, respectively. All control variables are at the 1% and 99% percentile levels, and all continuous variables are standardized afterwards. Firm-specific effects and year-specific effects are included. Standard errors are presented in parentheses. Finally, *, ** and *** represent significance levels of 10%, 5% and 1%, respectively.

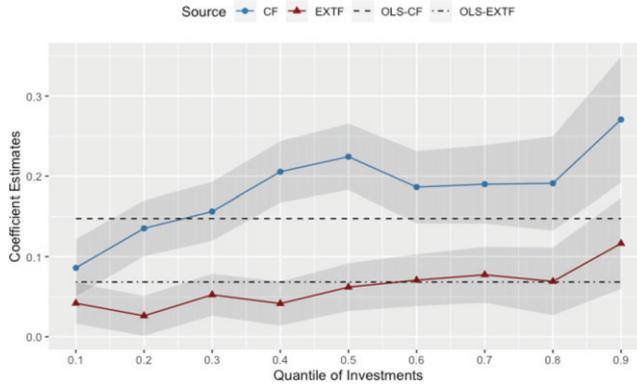


Figure 7. First rung of pecking order.

To take a further step, external funds (EXTF) is divided into two separate components, i.e., debt issues and equity issues. This derivative equation allows us to test the second rung of the pecking order theory:

$$q_{\tau}(I_{it}) = \alpha_{i,\tau} + \delta_{1,\tau}CF_{it} + \delta_{2,\tau}DEBT_{it} + \delta_{3,\tau}EQUITY_{it} + \sum_{k=1}^K \gamma_{k,\tau}Control_{it,k}, \quad (13)$$

where the parameters of interest are $\delta_{1,\tau}$, $\delta_{2,\tau}$, and $\delta_{3,\tau}$, representing the responsiveness of investments to internal funds, debt issues, and equity issues, respectively.

4.2. Data and sample construction

To generate our sample, we begin with all listed companies in the U.S. that have annual data regarding financial information in *Compustat*. We construct a balanced panel of non financial and non utility companies by drawing the sample every five-years, i.e., 1996, 2001, 2006, 2011, and 2016. We exclude observations with missing values from the sample and trim the outliers in all key variables following the rules as those in Cleary (1999). Therefore, our final sample consists of 3785 firm-year observations in total with $N = 757$, the number of firms, and $T = 5$, the number of years.

4.3. Empirical results

4.3.1. Testing on the first rung of pecking order theory

Panel A in Table 2 provides the panel quantile regression results as well as the OLS regression results estimated for the first rung of the pecking order theory (i.e., internal to external funds). The OLS results illustrate a simple situation on average, while the results of the quantile regression model depict a full picture of the financing hierarchy across different scales of investments.

The results of OLS regression indicate that both internal and external funds have statistically significant effect on investment activities. Consistent with the assumption in the first rung of pecking order theory, the responsiveness of investments to internal funds (0.148) is significantly more profound than external funds (0.068).

Concerning the results of the panel quantile regression, the financing hierarchy of investment activities gets more complicated across different scales of investments. The estimates of $\delta_{1,\tau}$ and $\delta_{2,\tau}$ are positive and statistically significant with p -values < 0.1 at all levels of investments. To illustrate the trends of each estimate and compare the relative relationship between them, we then plot the estimates of internal and external funds as shown in Figure 7. The horizontal axis measures different scales of investment activities and the vertical axis measures the estimated coefficients of different sources of funds. The solid points indicate the point estimates from the panel quantile regression and the shaded

area represents the corresponding 90% confidence interval. The two dashed lines indicate the estimates of internal and external funds from the OLS regression, respectively. It is observed that as the investment quantile ranges from 10% to 90%, the magnitude of coefficients of external funds generally experiences an upward trend from 0.042 to 0.119. In the mean time, the magnitude of coefficients of internal funds rises to the second highest point at the 50% quantile and declines at the 60% quantile. Then the magnitude of coefficients of internal funds remain stable from 60% to 80% quantiles while it rises to the peak at the highest quantile. Generally, the estimated coefficients of internal funds are significantly larger than that of external funds except for the lowest investments quantile. Our results, thus, suggest that the corporate financing hierarchy is in accordance with the first rung of the pecking order theory across all investments quantiles from 10% to 90%. It is reasonable that our results supports the view that internal funds are preferred to external funds, as prior literature indicate that internal funds help firms avoid transaction costs and information asymmetry costs on debt or equity issuances (Faulkender and Wang, 2006) and liquidation cost on debt issuance (Al-Najjar and Belghitar, 2011).

4.3.2. Testing on the second rung of pecking order theory

Provided that managers in pursuit of maximizing the value for original shareholders have information advantages over outside creditors and investors, the second rung of the pecking order hypothesis theoretically assumes that managers should use debt and then equity issues in sequence to invest projects with a positive net present value. Accordingly, we take a further empirical test for the second rung of the pecking order theory.

Panel B in Table 2 represents the regression results estimated where external funds are divided into funds raised by debt issues and equity issues. The results, however, are inconsistent with the pecking order hypothesis in the second rung for both OLS regression and the panel quantile regression, despite that the first run still holds. For all quantiles of investments, the estimates of $\delta_{2,\tau}$ and $\delta_{3,\tau}$ are positive, and the estimates of $\delta_{2,\tau}$ are statistically significant with p -values < 0.1 . However, the estimates of $\delta_{3,\tau}$ are statistically significant only at 30% to 70% quantiles with p -values < 0.1 but not at two tailed quantiles.

To compare the magnitudes of estimated coefficients on cash flow, debt issues and equity issues and determine their relative importance for investment activities at each investment level, we plot the estimates of internal and two types of external funds with the similar patterns as Figure 7. As is shown in Figure 8, firms prefer internal funds to both debt and equity issues across all investment levels, in support of the first rung of the pecking order theory. As for the second rung, the magnitude of $\delta_{2,\tau}$ keeps an upward tendency while that $\delta_{3,\tau}$ generally experience growth but with slight fluctuations at some of the quantiles. However, the shaded areas for the estimates of debt issues and equity issues overlap for all quantiles of investments. It implies that the relative magnitudes of the two sources of external funds are not significantly distinctive from each other, resulting in an evidence against the second rung of the pecking order theory.

Although our results reveal some interesting patterns seemingly contradictory to the traditional pecking order hypothesis, we acknowledge that the results above are not against the view that information asymmetry is the driving force of adverse selection and then corporate financing behavior in terms of pecking order theory. Still, uninformed investors rationally demand different premiums on debt issues and equity issues given that firms are confronted with severe adverse selection problems (Myers, 1984; Myers and Majluf, 1984). Our results could be attributable to the improved information environments in the U.S. thanks to the gradually updated FASB standards (Khan et al., 2018) and the growing number of small listed companies since 1990s (Chay et al., 2015; Frank and Goyal, 2003). Frank and Goyal (2003) documented that the net equity issues track the financing deficit much more closely than net debt issues do, and the bulk of external financing takes the form of equity. A particularly striking phenomenon is that a lot more small and unprofitable firms with less severe adverse selection problems became publicly traded during the 1990s so that they do not behave according to pecking order theory.

The overall results indicate that the financing hierarchy departs from the pecking order theory at any of the investments quantiles, as presented clearly in Figure 4. Meanwhile, our results are

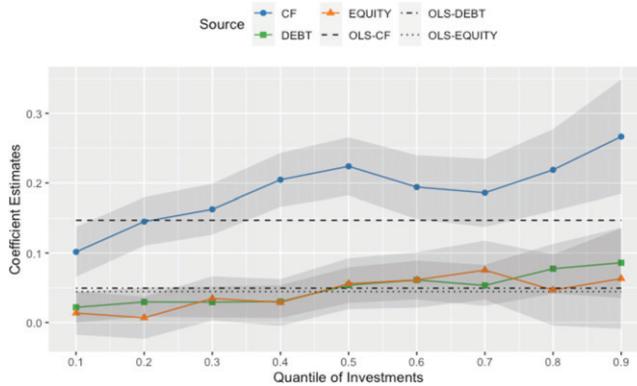


Figure 8. Second rung of pecking order.

distinctive from the results documented by previous studies and we presume the difference mainly arises from our deployment of the panel quantile regression controlling firm-specific and year fixed effects.

5. Conclusion

This article studies the estimation and inference of a quantile regression model for panel data with correlated random effects. Under the framework of Chamberlain (1982) and Abrevaya and Dahl (2008), we propose a new estimation method, which takes the influence of disturbance into account and allows the time period T to be fixed, by using the integrated quasi-likelihood function for conditional quantile model, and the Laplace approximation to relax the normality assumption. The Monte Carlo simulation results confirm the asymptotic theories of the proposed estimator and the asymptotic χ^2 test. An application to test the pecking order theory by using non financial and non utility U.S. firms in *Compustat* database has been investigated.

Although our method has many advantages over the previous literature, there still exist some limitations of our proposed method. First, the asymptotic theory of the estimation of disturbance variance σ^2 has not been established, which is left for further research. Second, the assumptions that $\{v_i\}$ are independent of $\{x_{it}\}$, and $\{y_{it}, x_{it}\}$ are i.i.d. across individual i , are relatively strong, which may restrict the application of our model.

In addition, we note several possible extensions of the present study. For example, it is worth pointing out that y_{it} might not be independent among individuals for many applications. To characterize the dependence among individuals for panel data, it might need to assume that T is large as discussed in Kato, Galvao, and Montes-Rojas (2012). For such a way, three methods of using factors are commonly used in the literature: the common correlated effect by Pesaran (2006), the interactive fixed effect approach proposed in Bai (2009), and the quantile factor model by Chen, Dolado, and Gonzalo (2021), so that our quantile model for panel data can be generalized to quantile model for panel data with cross-sectional dependence and large T . We leave such extensions as possible future research topics.

A. Assumptions and proofs

First, we make the following assumptions, which are necessary to establish the asymptotic normality of our estimator by following the analysis of Cai and Xu (2008) and Cai and Xiao (2012). All notations can be found in Section 2.3.

Assumptions:

- A1.** The series $\{y_{it}, \mathbf{x}_{it}\}$ is i.i.d. across individual i but can be correlated over t for fixed i ;
- A2.** The series $\{v_{i,\tau}\}$ is i.i.d. and independent of $\{\mathbf{x}_{it}\}$, and has a smooth and unimodal density function with zero mean;
- A3.** $\Sigma_x = E(\mathbf{x}_{it}\mathbf{x}_{it}^T)$ is invertible;
- A4.** $E(\|\mathbf{x}_{it}\|^\theta) < \infty$ with $\theta > 4$.

Assumption A1 assumes that the observations $\{\mathbf{x}_{it}, y_{it}\}$ are i.i.d. across individual i , but can be correlated between different time t for fixed i , which is similar to Assumption A1 in Kato, Galvao, and Montes-Rojas (2012). Assumption A2 is for model identification. For the non zero mean of $v_{i,\tau}$, one can always subtract the non zero mean and add it to the constant coefficient. Assumptions A3-A4 about the observations $\{\mathbf{x}_{it}\}$ guarantee the matrix $\mathbf{G}^T\mathbf{G}$ in model (3) is invertible, and $E(\|\mathbf{g}_{it}\|)^\theta < \infty$ for some $\theta > 4$

Next, we give the proof of Theorems 1 and 2. To this end, two lemmas and one proposition are needed as follows.

Lemma 1. Let $V_{NT}(\Delta)$ be a vector function that satisfies

- (i) $-\Delta^T V_{NT}(\lambda\Delta) \geq -\Delta^T V_{NT}(\Delta), \lambda \geq 1,$
- (ii) $\sup_{\|\Delta\| \leq M} \|V_{NT}(\Delta) - A_{NT} + D\Delta\| = o_p(1),$

where $\|A_{NT}\| = O_p(1), 0 < M < \infty,$ and D is a positive-definite matrix. Suppose $\hat{\Delta}$ is a vector such that $V_{NT}(\hat{\Delta}) = o_p(1).$ Then, we have $\|\hat{\Delta}\| = O_p(1)$ and

$$\hat{\Delta} = D^{-1}A_{NT} + o_p(1).$$

Proof. See the proof of Lemma A.4 in Koenker and Zhao (1996). □

Proposition 1. Some properties about $\psi_\tau(u, \sigma) = \tau - h_\tau(u, \sigma)$ are listed as follows.

- (i) $\psi_\tau(u, \sigma)$ is a non decreasing function about u when σ is fixed.
- (ii) $\psi_\tau(u, \sigma)$ is a bounded function, and both the first and second derivatives are also bounded.
- (iii) $\psi_\tau(u, \sigma)$ has a continuous second derivative function about u .

Proof. (i): Since $\psi_\tau(u, \sigma) = \tau - h_\tau(u, \sigma),$ we only need to show that $h_\tau(u, \sigma)$ is a non increasing function about parameter u when σ is fixed. Denote $m_\tau(u) \equiv h_\tau(u, \sigma),$ then

$$\begin{aligned} m_\tau(u) &= \frac{\partial \ln(v_\tau(u, \sigma))}{\partial u} \\ &= \frac{\exp\{u + \frac{1-2\tau}{2}\sigma^2\} \Phi(-\frac{u}{\sigma} + (\tau - 1)\sigma)}{\exp\{u + \frac{1-2\tau}{2}\sigma^2\} \Phi(-\frac{u}{\sigma} + (\tau - 1)\sigma) + \Phi(\frac{u}{\sigma} - \tau\sigma)} \equiv \frac{p(u)}{q(u)}, \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Since $m'_\tau(u) = [p'(u)q(u) - p(u)q'(u)]/q^2(u),$ we only need to show that for any $u \in \mathbb{R},$

$$\begin{aligned} m_1(u) &\equiv p'(u)q(u) - p(u)q'(u) \\ &= \exp\left\{u + \frac{1-2\tau}{2}\sigma^2\right\} \left\{ \Phi\left(\frac{u}{\sigma} - \tau\sigma\right) \left[\Phi\left(-\frac{u}{\sigma} + (\tau - 1)\sigma\right) - \frac{1}{\sigma}\phi\left(-\frac{u}{\sigma} + (\tau - 1)\sigma\right) \right] \right. \\ &\quad \left. - \frac{1}{\sigma}\Phi\left(-\frac{u}{\sigma} + (\tau - 1)\sigma\right) \phi\left(\frac{u}{\sigma} - \tau\sigma\right) \right\} \leq 0, \end{aligned} \tag{A.1}$$

where $\phi(\cdot)$ is the density probability function of the standard normal distribution. Let x denote $-u/\sigma + (\tau - 1)\sigma.$ Then

$$\begin{aligned} m_2(x) &\equiv m_1(u) / \exp\{u + (1 - 2\tau)\sigma^2/2\} \\ &= \Phi(-x - \sigma)\Phi(x) - \frac{1}{\sigma}\Phi(-x - \sigma)\phi(x) - \frac{1}{\sigma}\Phi(x)\phi(-x - \sigma). \end{aligned}$$

Notice that $m_2(x)$ is a symmetric function about $x = -\sigma/2$ and $\lim_{x \rightarrow \pm\infty} m_2(x) = 0$, so that (A.1) follows from the following fact that

$$m'_2(x) \geq 0, \quad x > -\sigma/2. \tag{A.2}$$

The next step is to show (A.2) holds. We only prove (A.2) when $\sigma = 1$. For $0 < \sigma < 1$ and $\sigma > 1$, one can show it similarly. For $\sigma = 1$, $m_2(x) = \Phi(-x-1)\Phi(x) - \Phi(-x-1)\phi(x) - \Phi(x)\phi(-x-1)$, and the first derivative of $m_2(x)$ is $m'_2(x) = \phi(x+1)\Phi(x)x + \Phi(-x-1)\phi(x)(1+x)$, which holds by the symmetry of $\phi(x)$. Obviously, $m'_2(-1/2) = 0$, $m'_2(0) > 0$, and $m'_2(x) \geq 0$ for $x \geq 0$. Next, we still needs to prove that

$$m'_2(x) \geq 0, \quad -\frac{1}{2} < x < 0. \tag{A.3}$$

For $-\frac{1}{2} < x < 0$, $-x^2 - x + 1 \geq 1$, the second derivative of $m_2(x)$ is

$$\begin{aligned} m''_2(x) &= (-x^2 - x + 1) [\phi(x+1)\Phi(x) + \Phi(-x-1)\phi(x)] - \phi(x+1)\phi(x) \\ &\geq \phi(x+1)\Phi(x) + \Phi(-x-1)\phi(x) - \phi(-x-1)\phi(x) \equiv m_3(x). \end{aligned}$$

By $m'_2(-\frac{1}{2}) = 0$ and $m'_2(0) > 0$, (A.3) is implied by

$$m_3(x) \geq 0, \quad -\frac{1}{2} < x < 0. \tag{A.4}$$

We now consider $m_3(x)$ for $-\frac{1}{2} < x < 0$. It is easy to show that

$$\begin{aligned} m_3(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+1)^2}{2}} \Phi(x) + \Phi(-x-1) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} - \phi(x+1) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ &= \left\{ e^{-\frac{2x+1}{2}} \Phi(x) + \Phi(-x-1) - \phi(x+1) \right\} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \end{aligned}$$

The first derivative of $\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right)^{-1} m_3(x)$ is

$$\begin{aligned} \left[\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right)^{-1} m_3(x) \right]' &= \left\{ -e^{-\frac{2x+1}{2}} \Phi(x) + e^{-\frac{2x+1}{2}} \phi(x) - \phi(x+1) - \phi(x+1)(-x-1) \right\} c_1 \\ &= \left\{ -e^{\frac{2x+1}{2}} \Phi(x) + \phi(x+1)(x+1) \right\} c_1 \\ &= \left\{ -e^{\frac{2x+1}{2}} \Phi(x) + \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+1)^2}{2}} (x+1) \right\} c_1 \\ &= \{-\Phi(x) + \phi(x)(x+1)\} c_2, \end{aligned}$$

where c_1 and c_2 are positive constant numbers. The second derivative of $c_2^{-1} m_3(x)$ is

$$[(c_2)^{-1} m_3(x)]'' = -\phi(x) + \phi(x)(-x)(x+1) + \phi(x) = -x(x+1)\phi(x).$$

Notice that $m''_3(x) \geq 0$ for $-\frac{1}{2} < x < 0$, and $m''_3(0) = 0$, $m''_3(-\frac{1}{2}) > 0$. Therefore, $m'_3(x)$ is an increasing function for $-\frac{1}{2} < x < 0$. Combining the fact that $m'_3(0) < 0$ and $m'_3(-\frac{1}{2}) < 0$, we know that $m_3(x)$ is a decreasing function for $-\frac{1}{2} < x < 0$. By $m_3(0) > 0$ and $m_3(-\frac{1}{2}) > 0$, we can obtain that (A.4) holds.

The proofs of (ii) and (iii) follow from the proof of (1) and omitted. This finishes the proof of the proposition. \square

Lemma 2. Under the assumptions of Theorem 1,

$$\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(u_{it}, \sigma) \mathbf{g}_{it} \right\| = O_p(1). \tag{A.5}$$

Proof. We rewrite (A.5) as follows

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \right) \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^N \Psi_i.$$

By Proposition 1, Ψ_i has a mean $\mu_{\Psi} = 0$ and a variance $V_{\Psi} < \infty$. And by Lindeberg-Lévy central limit theory,

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \xrightarrow{d} \mathcal{N}(0, V_{\Psi}).$$

Then, by Assumption A4 and Slutsky's theorem,

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it} \xrightarrow{d} \mathcal{N}(0, V_{\Psi} E(\mathbf{g}_{it}) E(\mathbf{g}_{it}^T))$$

Therefore,

$$\left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it} \right\| = O_p(1).$$

□

A.1. Proofs for Section 2.3

First, we derive the Bahadur representation of the panel quantile regression. We need to verify the conditions listed in Lemma 1. Define an objective function as follows.

$$Z_{NT}(\Delta) = \sum_{i=1}^N \sum_{t=1}^T \{ \ln l_{\tau}(u_{it} - \Delta^T b_{NT}^{-1} \mathbf{g}_{it}) - \ln l_{\tau}(u_{it}) \},$$

where $b_{NT} = \sqrt{NT}$, and $l_{\tau}(u)$ is actually $l_{\tau}(u, \sigma)$ that is defined in (7). In this proof, however, σ will be omitted for the estimation of ξ_{τ} is performed with some fixed σ . Indeed, $Z_{NT}(\Delta)$ is the objective function of the following model:

$$q_{\tau}(u_{it}) = (\Delta b_{NT}^{-1})_{\tau}^T \mathbf{g}_{it},$$

and the QMLE of $(\Delta b_{NT}^{-1})_{\tau}$ is $(\hat{\xi}_{\tau} - \xi_{\tau})$ by (8). Therefore, we can obtain

$$\frac{\partial Z_{NT}(\Delta)}{\partial \Delta} \Big|_{\Delta = \sqrt{NT}(\hat{\xi}_{\tau} - \xi_{\tau})} = 0. \tag{A.6}$$

Using the identity which has a similar form in Knight (1998), we have

$$-\ln l_{\tau}(u - v) + \ln l_{\tau}(u) = -v \psi_{\tau}(u) + \int_0^v [h(u - s) - h(u)] ds,$$

where $\psi_{\tau}(u) = -\partial \ln l_{\tau}(u) / \partial u$. Then,

$$Z_{NT}(\Delta) = \sum_{i=1}^N \sum_{t=1}^T \Delta^T b_{NT}^{-1} \mathbf{g}_{it} \psi_{\tau}(u_{it}) - \sum_{i=1}^N \sum_{t=1}^T \int_0^{\Delta^T b_{NT}^{-1} \mathbf{g}_{it}} [h(u_{it} - s) - h(u_{it})] ds \tag{A.7}$$

By (A.7), $Z_{NT}(\Delta)$ is derivable. Then we define a new function as follows.

$$\begin{aligned} V_{NT}(\Delta) &\equiv \frac{\partial Z_{NT}(\Delta)}{\partial \Delta} \\ &= \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1} \mathbf{g}_{it} \psi_{\tau}(u_{it}) - \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1} \mathbf{g}_{it} [h(u_{it} - \Delta^T b_{NT}^{-1} \mathbf{g}_{it}) - h(u_{it})]. \end{aligned}$$

The next step is to show that $V_{NT}(\Delta)$ satisfies condition (i) in Lemma 1. To this end, it follows by Proposition 1 that $\psi_\tau(u)$ is a non decreasing function, so that

$$-\Delta^\top V_{NT}(\lambda\Delta) = \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(u_{it} - \lambda\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})(-\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})$$

is a non decreasing function of λ if $\Delta^\top b_{NT}^{-1}\mathbf{g}_{it} > 0$. Similarly,

$$-\Delta^\top V_{NT}(\lambda\Delta) = \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(u_{it} - \lambda\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})(-\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})$$

is a non decreasing function of λ if $\Delta^\top b_{NT}^{-1}\mathbf{g}_{it} < 0$. Thus,

$$-\Delta^\top V_{NT}(\lambda\Delta) = \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(u_{it} - \lambda\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})(-\Delta^\top b_{NT}^{-1}\mathbf{g}_{it})$$

is a non decreasing function of λ . Therefore, we have for $\lambda > 1$

$$-\Delta^\top V_{NT}(\lambda\Delta) \geq \Delta^\top V_{NT}(\Delta),$$

so that condition (i) of Lemma 1 is verified.

It still needs to prove that $V_{NT}(\Delta)$ satisfies condition (ii) in Lemma 1. To this end,

$$\begin{aligned} V_{NT}(\Delta) &= \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} \psi_\tau(u_{it}) - \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} [h(u_{it} - \Delta^\top b_{NT}^{-1}\mathbf{g}_{it}) - h(u_{it})] \\ &= \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} \psi_\tau(u_{it}) - \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it}|\mathbf{g}_{it}) - \sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E(\lambda_{it}|\mathbf{g}_{it})], \\ &= A_{NT} - \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it}|\mathbf{g}_{it}) - \sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E(\lambda_{it}|\mathbf{g}_{it})], \end{aligned} \quad (\text{A.8})$$

where $A_{NT} = \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} \psi_\tau(u_{it})$ and $\lambda_{it} = b_{NT}^{-1}\mathbf{g}_{it} [h(u_{it} - \Delta^\top b_{NT}^{-1}\mathbf{g}_{it}) - h(u_{it})]$. By Taylor's expansion,

$$\begin{aligned} \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it}|\mathbf{g}_{it}) &\equiv \sum_{i=1}^N \sum_{t=1}^T E_{it}(\lambda_{it}) \\ &= \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} E_{it} (h(u_{it} - \Delta^\top b_{NT}^{-1}\mathbf{g}_{it}) - h(u_{it})) \\ &= \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} E_{it} \left(-\dot{h}(u_{it})\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta + \frac{1}{2}\ddot{h}(u_{it}^*)\Delta^\top b_{NT}^{-1}\mathbf{g}_{it}\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta \right) \\ &= \sum_{i=1}^N \sum_{t=1}^T E(-\dot{h}(u_{it})|\mathbf{g}_{it}) b_{NT}^{-1}\mathbf{g}_{it}\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta + \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1}\mathbf{g}_{it} \frac{1}{2} E(\ddot{h}(u_{it}^*)|\mathbf{g}_{it}) \Delta^\top b_{NT}^{-1}\mathbf{g}_{it}\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta \\ &\equiv B_1 + B_2, \end{aligned} \quad (\text{A.9})$$

where $E_{it}(\cdot)$ denotes the conditional expectation dependent on \mathbf{g}_{it} and u_{it}^* is the interior point between u_{it} and $u_{it} + \Delta^\top b_{NT}^{-1}\Delta\mathbf{g}_{it}$. Note that $\dot{h} < 0$ by Proposition 1(i). We decompose B_1 into two parts as follows.

$$\begin{aligned} B_1 &= \sum_{i=1}^N \sum_{t=1}^T E(-\dot{h}(u_{it})) b_{NT}^{-1}\mathbf{g}_{it}\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta \\ &\quad + \sum_{i=1}^N \sum_{t=1}^T \{E(-\dot{h}(u_{it})|\mathbf{g}_{it}) - E[E(-\dot{h}(u_{it})|\mathbf{g}_{it})]\} b_{NT}^{-1}\mathbf{g}_{it}\mathbf{g}_{it}^\top b_{NT}^{-1}\Delta, \end{aligned}$$

and rewrite the second part as follows

$$\begin{aligned} & \sum_{i=1}^N \sum_{t=1}^T \{E(-\dot{h}(u_{it})|\mathbf{g}_{it}) - E[E(-\dot{h}(u_{it})|\mathbf{g}_{it})]\} b_{NT}^{-1} \mathbf{g}_{it} \mathbf{g}_{it}^T b_{NT}^{-1} \\ &= \frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \{E(-\dot{h}(u_{it})|\mathbf{g}_{it}) - E[E(-\dot{h}(u_{it})|\mathbf{g}_{it})]\} \mathbf{g}_{it} \mathbf{g}_{it}^T \cdot \frac{1}{\sqrt{NT}}, \end{aligned}$$

which will be shown to be $o_p(1)$. To do this, because

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \sum_{t=1}^T \{E(-\dot{h}(u_{it})|\mathbf{g}_{it}) - E[E(-\dot{h}(u_{it})|\mathbf{g}_{it})]\} \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

where $\Sigma = \text{Var}[E(-\dot{h}(u_{it})|\mathbf{g}_{it})] E(\mathbf{g}_{it} \mathbf{g}_{it}^T) E(\mathbf{g}_{it} \mathbf{g}_{it}^T)$ by Proposition 1(ii) that $h(u)$ is bounded and Assumption A4, it follows that

$$\sum_{i=1}^N \sum_{t=1}^T \{E(-\dot{h}(u_{it})|\mathbf{g}_{it}) - E[E(-\dot{h}(u_{it})|\mathbf{g}_{it})]\} b_{NT}^{-1} \mathbf{g}_{it} \mathbf{g}_{it}^T b_{NT}^{-1} = O_p(1) \cdot o(1) = o_p(1).$$

Therefore,

$$B_1 = E[\mathbf{g}_{it} \mathbf{g}_{it}^T E(-\dot{h}(u_{it}, \sigma))] \Delta + o_p(1) \equiv D_{g\dot{h}}(\sigma) \Delta + o_p(1), \tag{A.10}$$

where $D_{g\dot{h}}(\sigma) = E[\mathbf{g}_{it} \mathbf{g}_{it}^T E(-\dot{h}(u_{it}, \sigma))]$. Also,

$$\begin{aligned} \|B_2\| &= \left\| \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1} \mathbf{g}_{it} \frac{1}{2} E(\ddot{h}(u^*)|\mathbf{g}_{it}) \Delta^T b_{NT}^{-1} \mathbf{g}_{it} \mathbf{g}_{it}^T b_{NT}^{-1} \Delta \right\| \\ &\leq b_{NT}^{-1} \frac{1}{2} |E(\ddot{h}(u^*)|\mathbf{g}_{it})| \cdot \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-2} \|\mathbf{g}_{it}\|^3 \|\Delta\|^2 \\ &\xrightarrow{p} o(1) \cdot \frac{1}{2} |E(\ddot{h}(u^*)|\mathbf{g}_{it})| E\|\mathbf{g}_{it}\|^3 \|\Delta\|^2 = 0, \end{aligned}$$

by Assumption A4, $\|\Delta\| \leq M$ in Lemma 1, and Proposition 1(ii). Thus,

$$B_2 = o_p(1). \tag{A.11}$$

By combining (A.10) and (A.11), we have

$$\sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it}|\mathbf{g}_{it}) = D_{g\lambda}(\sigma) + o_p(1). \tag{A.12}$$

Next, we will verify that

$$\sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E(\lambda_{it}|\mathbf{g}_{it})] = o_p(1).$$

To this end, we can show that for $\Delta^\top b_{NT}^{-1} \mathbf{g}_{it} > 0$, $h(u_{it} + \Delta^\top b_{NT}^{-1} \mathbf{g}_{it}) - h(u_{it}) > 0$, and $[h(u_{it} + \Delta^\top b_{NT}^{-1} \mathbf{g}_{it}) - h(u_{it})] \in (0, 1)$ by $h(u) \in (0, 1)$,

$$\begin{aligned} \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it} \lambda_{it}^\top) &= \sum_{i=1}^N \sum_{t=1}^T E[E(\lambda_{it} \lambda_{it}^\top | \mathbf{g}_{it})] \\ &= \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T E \left\{ E[(h(u_{it} + \Delta^\top b_{NT}^{-1} \mathbf{g}_{it}) - h(u_{it}))^2 \mathbf{g}_{it} \mathbf{g}_{it}^\top | \mathbf{g}_{it}] \right\} \\ &\leq E \left\{ E[(h(u_{it} + \Delta^\top b_{NT}^{-1} \mathbf{g}_{it}) - h(u_{it})) \mathbf{g}_{it} \mathbf{g}_{it}^\top | \mathbf{g}_{it}] \right\} \\ &= E \left\{ \dot{h}(u_{it}^*) \Delta^\top b_{NT}^{-1} \mathbf{g}_{it} \mathbf{g}_{it} \mathbf{g}_{it}^\top \right\} \\ &\leq \max_{u \in \mathbb{R}} |\dot{h}(u)| \cdot \frac{1}{\sqrt{NT}} \cdot E(\Delta^\top \mathbf{g}_{it} \mathbf{g}_{it} \mathbf{g}_{it}^\top), \end{aligned}$$

where $u_{it}^* \in (u_{it}, u_{it} + \Delta^\top b_{NT}^{-1} \mathbf{g}_{it})$. For those satisfying $\Delta^\top b_{NT}^{-1} \mathbf{g}_{it} < 0$, we can obtain the same result. Thus, for $0 < M < \infty$,

$$\left\| \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it} \lambda_{it}^\top) \right\| \leq \max_{u \in \mathbb{R}} |\dot{h}(u)| \cdot \frac{1}{\sqrt{NT}} \cdot M \cdot E \|\mathbf{g}_{it}\|^3 = o(1),$$

by Proposition 1(ii) and Assumption A4. As a result, we prove that

$$\sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it} \lambda_{it}^\top) = o(1).$$

Therefore,

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E(\lambda_{it} | \mathbf{g}_{it})] \right) &\equiv \text{Var} \left(\sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E_{it}(\lambda_{it})] \right) \\ &= \sum_{i=1}^N \sum_{t=1}^T \text{Var}[\lambda_{it} - E_{it}(\lambda_{it})] + \sum_{i=1}^N \sum_{t \neq s} \text{Cov}[\lambda_{it} - E_{it}(\lambda_{it}), \lambda_{is} - E_{is}(\lambda_{is})] \\ &\leq \sum_{i=1}^N \sum_{t=1}^T \text{Var}[\lambda_{it} - E_{it}(\lambda_{it})] + (T-1) \sum_{i=1}^N \sum_{t=1}^T \text{Var}[\lambda_{it} - E_{it}(\lambda_{it})] \\ &= T \sum_{i=1}^N \sum_{t=1}^T \text{Var}[\lambda_{it} - E_{it}(\lambda_{it})] = T \sum_{i=1}^N \sum_{t=1}^T E[\lambda_{it} - E_{it}(\lambda_{it})][\lambda_{it} - E_{it}(\lambda_{it})]^\top \\ &\leq T \sum_{i=1}^N \sum_{t=1}^T E(\lambda_{it} \lambda_{it}^\top) = o(1). \end{aligned}$$

Thus, by Chebyshev's inequality, it concludes that

$$\sum_{i=1}^N \sum_{t=1}^T [\lambda_{it} - E(\lambda_{it} | \mathbf{g}_{it})] = o_p(1). \quad (\text{A.13})$$

By (A.8), (A.12), and (A.13), for $\|\Delta\| \leq M$, $0 < M < \infty$,

$$V_{NT}(\Delta) = A_{NT} - D_{gh}(\sigma) + o_p(1)$$

Then, for $0 < M < \infty$,

$$\sup_{\|\Delta\| \leq M} \left\| V_{NT}(\Delta) + D_{gh}(\sigma) - A_{NT} \right\| = o_p(1). \quad (\text{A.14})$$

By Lemma 2,

$$\|A_{NT}\| = \left\| \sum_{i=1}^N \sum_{t=1}^T b_{NT}^{-1} \mathbf{g}_{it} \psi_{\tau}(u_{it}) \right\| = O_p(1), \tag{A.15}$$

and by (A.6),

$$\left| V_{NT} \left(\sqrt{NT}(\hat{\xi}_{\tau} - \xi_{\tau}) \right) \right| = \left| \frac{\partial Z_{NT}(\Delta)}{\partial \Delta} \Big|_{\Delta = \sqrt{NT}(\hat{\xi}_{\tau} - \xi_{\tau})} \right| = 0 = o_p(1). \tag{A.16}$$

Then by (A.14), (A.15), and (A.16), we conclude that condition (ii) of Lemma 1 is verified. Thus, we obtain the Bahadur representation

$$\sqrt{NT}(\hat{\xi}_{\tau} - \xi_{\tau}) \equiv \hat{\Delta} = D_{gh}^{-1}(\sigma)A_{NT} + o_p(1). \tag{A.17}$$

A.1.1. Completion proof of Theorem 1

From (A.17), (A.15), and the fact that T is fixed, clearly we know that $\sqrt{N}\|\hat{\xi}_{\tau} - \xi_{\tau}\| = O_p(1)$, and thus $\|\hat{\xi}_{\tau} - \xi_{\tau}\| = O_p(1/\sqrt{N})$. Since $\delta_{\tau} = \mathbf{e}_1^{\top} \xi_{\tau}$ and $\hat{\delta}_{\tau} = \mathbf{e}_1^{\top} \hat{\xi}_{\tau}$, we have

$$\|\hat{\delta}_{\tau} - \delta_{\tau}\| = O_p(1/\sqrt{N}) = o_p(1),$$

which implies that our proposed estimator $\hat{\delta}_{\tau}$ converges to δ_{τ} in probability with \sqrt{N} -rate.

A.1.2. Completion proof of Theorem 2

Clearly, the Bahadur representation of $\hat{\delta}_{\tau}$ is

$$\sqrt{N}(\hat{\delta}_{\tau} - \delta_{\tau}) \simeq \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{e}_1^{\top} D_{gh}^{-1}(\sigma) \frac{1}{T} \sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^N \Gamma_i.$$

It is easy to see the mean of Γ_i is zero for $E[\psi_{\tau}(u, \sigma)] = 0$ and the variance of Γ_i is given by

$$\begin{aligned} \text{Var}(\Gamma_i) &= \mathbf{e}_1^{\top} D_{gh}^{-1}(\sigma) \frac{1}{T^2} \text{Var} \left(\sum_{t=1}^T \psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it} \right) D_{gh}^{-1}(\sigma) \mathbf{e}_1 \\ &= \mathbf{e}_1^{\top} D_{gh}^{-1}(\sigma) \left\{ \text{Var}[\psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it}] + \sum_{t=2}^T \frac{2(T-t+1)}{T^2} \text{Cov}(\psi_{\tau}(u_{i1}, \sigma) \mathbf{g}_{i1}, \psi_{\tau}(u_{it}, \sigma) \mathbf{g}_{it}) \right\} \\ &\quad \times D_{gh}^{-1}(\sigma) \mathbf{e}_1 \\ &= \mathbf{e}_1^{\top} D_{gh}^{-1}(\sigma) \left\{ E[\psi_{\tau}(u_{it}, \sigma)]^2 E[\mathbf{g}_{it} \mathbf{g}_{it}^{\top}] + \sum_{t=2}^T \frac{2(T-t+1)}{T^2} E[\psi_{\tau}(u_{i1}, \sigma) \psi_{\tau}(u_{it}, \sigma)] E[\mathbf{g}_{i1} \mathbf{g}_{it}^{\top}] \right\} \\ &\quad \times D_{gh}^{-1}(\sigma) \mathbf{e}_1. \end{aligned}$$

Finally, since Γ_i is i.i.d. across individual i , we can follow the Lindeberg-Lévy central limit theorem to get the asymptotic normality of $\hat{\delta}_{\tau}$.

Acknowledgments

The authors express their sincere thanks to the editor, Professor Yuya Saaki, the associate editor, and three anonymous referees for their great comments and suggestions that improved significantly the quality and the presentation of the article. Also, they thank Mr. Zihan Wang for his help to the revision.

Disclosure Statement

The authors claim that there are no relevant financial or non financial competing interests to report for this article. Also, the authors declare that they do not use any generative AI and AI assisted technologies in the writing process.

Funding

This research was partially supported from the National Natural Science Foundation of China (NSFC) with grant numbers 72133002, 72033008 and 72374201.

References

- Abrevaya, J., Dahl, C. M. (2008). The effects of birth inputs on birthweight. *Journal of Business & Economic Statistics* 26(4):379–397. doi: [10.1198/073500107000000269](https://doi.org/10.1198/073500107000000269)
- Al-Najjar, B., Belghitar, Y. (2011). Corporate cash holdings and dividend payments: evidence from simultaneous analysis. *Managerial and Decision Economics* 32(4):231–241. doi: [10.1002/mde.1529](https://doi.org/10.1002/mde.1529)
- Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica* 77(4):1229–1279.
- Bharath, S. T., Pasquariello, P., Wu, G. (2009). Does asymmetric information drive capital structure decisions?. *Review of Financial Studies* 22(8):3211–3243. doi: [10.1093/rfs/hhn076](https://doi.org/10.1093/rfs/hhn076)
- Brennan, M., Kraus, A. (1987). Efficient financing under asymmetric information. *The Journal of Finance* 42(5):1225–1243. doi: [10.1111/j.1540-6261.1987.tb04363.x](https://doi.org/10.1111/j.1540-6261.1987.tb04363.x)
- Breslow, N. E., Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88(421):9–25. doi: [10.1080/01621459.1993.10594284](https://doi.org/10.1080/01621459.1993.10594284)
- Brown, J. R., Fazzari, S. M., Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *The Journal of Finance* 64(1):151–185. doi: [10.1111/j.1540-6261.2008.01431.x](https://doi.org/10.1111/j.1540-6261.2008.01431.x)
- Brown, S., Dutordoir, M., Veld, C., Veld-Merkoulova, Y. (2019). What is the role of institutional investors in corporate capital structure decisions? a survey analysis. *Journal of Corporate Finance* 58:270–286. doi: [10.1016/j.jcorpfin.2019.05.001](https://doi.org/10.1016/j.jcorpfin.2019.05.001)
- Cai, Z., Chen, L., Fang, Y. (2018). A semiparametric quantile panel data model with an application to estimating the growth effect of FDI. *Journal of Econometrics* 206(2):531–553. doi: [10.1016/j.jeconom.2018.06.013](https://doi.org/10.1016/j.jeconom.2018.06.013)
- Cai, Z., Xiao, Z. (2012). Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. *Journal of Econometrics* 167(2):413–425. doi: [10.1016/j.jeconom.2011.09.025](https://doi.org/10.1016/j.jeconom.2011.09.025)
- Cai, Z., Xu, X. (2008). Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association* 103(484):1595–1608. doi: [10.1198/016214508000000977](https://doi.org/10.1198/016214508000000977)
- Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics* 18(1):5–46. doi: [10.1016/0304-4076\(82\)90094-X](https://doi.org/10.1016/0304-4076(82)90094-X)
- Chamberlain, G. (1984). Panel Data. In *Handbook of Econometrics*, Vol. 2. Amsterdam: North-Holland.
- Chay, J. B., Park, S. H., Kim, S., Suh, J. (2015). Financing hierarchy: Evidence from quantile regression. *Journal of Corporate Finance* 33:147–163. doi: [10.1016/j.jcorpfin.2015.06.004](https://doi.org/10.1016/j.jcorpfin.2015.06.004)
- Chen, L., Dolado, J. J., Gonzalo, J. (2021). Quantile factor models. *Econometrica* 89(2):875–910. doi: [10.3982/ECTA15746](https://doi.org/10.3982/ECTA15746)
- Cleary, S. (1999). The relationship between firm investment and financial status. *The Journal of Finance* 54(2):673–692. doi: [10.1111/0022-1082.00121](https://doi.org/10.1111/0022-1082.00121)
- Fama, E. F., French, K. R. (2002). Testing trade-off and pecking order predictions about dividends and debt. *Review of Financial Studies* 15(1):1–33. doi: [10.1093/rfs/15.1.1](https://doi.org/10.1093/rfs/15.1.1)
- Fama, E. F., French, K. R. (2005). Financing decisions: who issues stock?. *Journal of Financial Economics* 76(3):549–582. doi: [10.1016/j.jfineco.2004.10.003](https://doi.org/10.1016/j.jfineco.2004.10.003)
- Faulkender, M., Wang, R. (2006). Corporate financial policy and the value of cash. *The Journal of Finance* 61(4):1957–1990. doi: [10.1111/j.1540-6261.2006.00894.x](https://doi.org/10.1111/j.1540-6261.2006.00894.x)
- Frank, M. Z., Goyal, V. K. (2003). Testing the pecking order theory of capital structure. *Journal of Financial Economics* 67(2):217–248. doi: [10.1016/S0304-405X\(02\)00252-0](https://doi.org/10.1016/S0304-405X(02)00252-0)
- Galvao, A. F., Lamarche, C., Lima, L. R. (2013). Estimation of censored quantile regression for panel data with fixed effects. *Journal of the American Statistical Association* 108(503):1075–1089. doi: [10.1080/01621459.2013.818002](https://doi.org/10.1080/01621459.2013.818002)
- Gu, J., Volgushev, S. (2019). Panel data quantile regression with grouped fixed effects. *Journal of Econometrics* 213(1):68–91. doi: [10.1016/j.jeconom.2019.04.006](https://doi.org/10.1016/j.jeconom.2019.04.006)
- Harding, M. C., Lamarche, C. (2017). Penalized quantile regression with semiparametric correlated effects: An application with heterogeneous preferences. *Journal of Applied Econometrics* 32(2):342–358. doi: [10.1002/jae.2520](https://doi.org/10.1002/jae.2520)
- Harding, M. C., Hausman, J. (2007). Using a laplace approximation to estimate the random coefficients logit model by nonlinear least squares*. *International Economic Review* 48(4):1311–1328. doi: [10.1111/j.1468-2354.2007.00463.x](https://doi.org/10.1111/j.1468-2354.2007.00463.x)

- Hausman, J., Liu, H., Luo, Y.E., Palmer, C. (2021). Errors in the dependent variable of quantile regression models. *Econometrica* 89(2):849–873. doi: [10.3982/ECTA14667](https://doi.org/10.3982/ECTA14667)
- Jung, K., Kim, Y.-C., Stulz, R. (1996). Timing, investment opportunities, managerial discretion, and the security issue decision. *Journal of Financial Economics* 42(2):159–185. doi: [10.1016/0304-405X\(96\)00881-1](https://doi.org/10.1016/0304-405X(96)00881-1)
- Kato, K., F. Galvao, A., Montes-Rojas, G. V. (2012). Asymptotics for panel quantile regression models with individual effects. *Journal of Econometrics* 170(1):76–91. doi: [10.1016/j.jeconom.2012.02.007](https://doi.org/10.1016/j.jeconom.2012.02.007)
- Khan, U., Li, B., Rajgopal, S., Venkatachalam, M. (2018). Do the fasb's standards add shareholder value?. *The Accounting Review* 93(2):209–247. doi: [10.2308/accr-51840](https://doi.org/10.2308/accr-51840)
- Knight, K. (1998). Limiting distributions for L1 regression estimators under general conditions. *Annals of Statistics* 26(2):755–770.
- Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis* 91(1):74–89. doi: [10.1016/j.jmva.2004.05.006](https://doi.org/10.1016/j.jmva.2004.05.006)
- Koenker, R., Bassett, G. (1978). Regression quantiles. *Econometrica* 46(1):33. doi: [10.2307/1913643](https://doi.org/10.2307/1913643)
- Koenker, R., Zhao, Q. (1996). Conditional quantile estimation and inference for arch models. *Econometric Theory* 12(5):793–813. doi: [10.1017/S0266466600007167](https://doi.org/10.1017/S0266466600007167)
- Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *Journal of Econometrics* 128(1):137–164. doi: [10.1016/j.jeconom.2004.08.010](https://doi.org/10.1016/j.jeconom.2004.08.010)
- Lamarche, C. (2010). Robust penalized quantile regression estimation for panel data. *Journal of Econometrics* 157(2):396–408. doi: [10.1016/j.jeconom.2010.03.042](https://doi.org/10.1016/j.jeconom.2010.03.042)
- Lemmon, M. L., Zender, J. F. (2010). Debt capacity and tests of capital structure theories. *Journal of Financial and Quantitative Analysis* 45(5):1161–1187. doi: [10.1017/S0022109010000499](https://doi.org/10.1017/S0022109010000499)
- Machado, J. A., Santos Silva, J. M. C. (2019). Quantiles via moments. *Journal of Econometrics* 213(1):145–173. doi: [10.1016/j.jeconom.2019.04.009](https://doi.org/10.1016/j.jeconom.2019.04.009)
- Myers, S. C. (1984). The capital structure puzzle. *The Journal of Finance* 39(3):575. doi: [10.2307/2327916](https://doi.org/10.2307/2327916)
- Myers, S. C., Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics* 13(2):187–221. doi: [10.1016/0304-405X\(84\)90023-0](https://doi.org/10.1016/0304-405X(84)90023-0)
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74(4):967–1012. doi: [10.1111/j.1468-0262.2006.00692.x](https://doi.org/10.1111/j.1468-0262.2006.00692.x)
- Powell, D. (2022). Quantile regression with nonadditive fixed effects. *Empirical Economics* 63(5):2675–2691. doi: [10.1007/s00181-022-02216-6](https://doi.org/10.1007/s00181-022-02216-6)
- Shyam-Sunder, L., C. Myers, S. (1999). Testing static tradeoff against pecking order models of capital structure. *Journal of Financial Economics* 51(2):219–244. doi: [10.1016/S0304-405X\(98\)00051-8](https://doi.org/10.1016/S0304-405X(98)00051-8)