Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Quantile Regression Models*

Matthew Harding†, Carlos Lamarche‡, and M. Hashem Pesaran§

January 19, 2017

Abstract

This paper proposes a quantile regression estimator for a heterogeneous panel model with lagged dependent variables and interactive effects. The paper adopts the Common Correlated Effects (CCE) approach proposed by Pesaran (2006) and demonstrates that the extension to the estimation of dynamic quantile regression models is feasible under similar conditions to the ones used in the literature. We establish consistency and derive the asymptotic distribution of the new quantile estimator. Monte Carlo studies are carried out to study the small sample behavior of the proposed approach. The evidence shows that the estimator can significantly improve the performance of existing estimators as long as the time series dimension of the panel is large. We present an application to the evaluation of Time-of-Use pricing using a large randomized control trial.

JEL: C21, C23, C25, C55.

Keywords: Common Correlated Effects; Dynamic Panel; Quantile Regression; Smart Meter; Randomized Experiment

*The authors would like to thank ...

†Department of Economics, University of California Irvine, SSPB 3207, Irvine, CA 9269; Email: harding1@uci.edu

‡Department of Economics, University of Kentucky, 223G Gatton College of Business and Economics, Lexington, KY 40506-0034; Phone: (859) 257 3371; Email: clamarche@uky.edu

§Department of Economics, University of Southern California, Kaprielian Hall 300, Los Angeles, CA 90089-0253; Email: pesaran@dornsife.usc.edu

1. Introduction

In the last decade, the literature on linear panel data models has made significant progress on the estimation of models with multi-factor error structure. Recent papers have focused on the estimation of models with a fixed number of unobserved factors (see e.g. Pesaran (2006), Bai (2009), Pesaran and Chudik (2014), Chudik and Pesaran (2015)). The Common Correlated Effects (CCE) approach of Pesaran (2006) is robust to cross-sectional dependence and slope heterogeneity. Chudik and Pesaran (2015) extend the approach developed by Pesaran (2006) to heterogeneous dynamic panel data models, for situations where the cross-sectional dimension (N) and the time-series dimension (T) are relatively large. More recently, Moon and Weidner (2015a, 2015b) develop estimation approaches for models with lagged dependent variables and cross-sectional dependence. These methods however are designed for Gaussian conditions and do not offer the possibility of estimating heterogeneous distributional effects, which is an important consideration for practice.

A number of different approaches have been developed for the estimation of a panel quantile regression model, offering alternatives to Gaussian models. The recent quantile regression literature include work by Koenker (2004), Lamarche (2010), Galvao (2011), Rosen (2012), Chernozhukov, Fernandez-Val, Hahn and Newey (2013) and Chernozhukov, Fernandez-Val, Hoderlein, Holzmann and Newey (2015), Harding and Lamarche (2014, 2016), Arellano and Bonhomme (2016), among others. Slope heterogeneity in quantile regression is investigated in Galvao and Wang (2015). With the exception of Galvao (2011) and Arellano and Bonhomme (2016), the literature has focused on estimating static models. Moreover, the quantile regression literature does not address cross-sectional dependence with the exception of Harding and Lamarche (2014) that adopt the approach proposed by Pesaran (2006) to estimate a static model with interactive effects. This paper extends the panel quantile literature to dynamic models with multi-factor error structure when both T and N are large.

We adopt a CCE approach and focus on estimation and inference of mean quantile coefficients. We allow for the possibility of endogenous independent variables and we study the conditions under which the slope parameter estimator is consistent. An important condition is that one plus the number of cross-sectional averages based on the independent variables be larger than the number of unobserved factors. Another important condition, which is similar to a condition used in Chudik and Pesaran (2015), is that the number of cross-sectional averages used to approximate the factors needs to be restricted to avoid inconsistent estimation of quantile coefficients. Under standard regularity conditions including T tending to infinity at a faster rate than N as in Kato, Galvao and Montes-Rojas (2012), we show that the average quantile estimator is consistent and asymptotically Gaussian. Moreover, we investigate the finite sample performance of the proposed approach in comparison with a candidate method for dynamic models. Using a comprehensive set of Monte Carlo experiments, we find that the proposed estimator has a satisfactory performance under different dynamic specifications when T is relatively large.

We apply the method to estimate how consumers respond to time-of-use (TOU) electricity pricing and different type of technologies that allow communication between customers and utility companies. In the empirical section, the quantile-specific demand equation allows us to estimate the short and long run impact of different enabling technologies, while including three key features of the problem: dynamics, slope heterogeneity and cross-sectional dependence. We use a data set of more than 6.5 million observations obtained from a large randomized control trial which includes $N = 779$ customers observed $T = 8639$ times.

Our findings suggest that smart thermostats are particularly effective relative to other enabling technologies and the differential effects are more pronounced at the lower tail of the conditional distribution of energy consumption. Smart thermostats, in addition of providing real time information on consumption and pricing, allow households to respond to price changes in advance by programming temperature settings for different times of the day. We also find that treated households appear to reduce overall consumption as a result of these technologies relative to the control group, but the average response does not summarize the distributional effect of the technologies. We also investigate the long-run effect of a change in energy price for different enabling technologies and we concentrate on the effects for different age and income groups.

The paper is organized as follows. The next section introduces the model and the proposed estimator. It also shows the asymptotic properties of the estimator. In Section 3, we offer simulation experiments to investigate the small sample performance of the proposed approach. Section 4 demonstrates how the estimator can be used in practice by exploring an application of electricity pricing and smart technology. Section 5 concludes. Mathematical proofs are provided in the Appendix and additional Monte Carlo results are offered in a Supplement.

2. Model and Assumptions

Consider the following panel data model for $i = 1, \ldots, N$ and $t = 1, \ldots, T$,

$$y_{it} = \alpha_i + \lambda_i y_{it-1} + \beta_i' \mathbf{x}_{it} + \varepsilon_{it}, \tag{2.1}$$

$$\varepsilon_{it} = \gamma_i' \mathbf{f}_t + \sigma(\mathbf{x}_{it}, \alpha_i, \mathbf{f}_t) u_{it}, \tag{2.2}$$

where $y_{it} \in \mathbb{R}$ is the response variable for cross-sectional i at time t, y_{it-1} is a lagged dependent variable, and,

$$\mathbf{x}_{it} = \mu_i + \mathbf{\Gamma}_i' \mathbf{f}_t + \mathbf{v}_{it}. \tag{2.3}$$

The variable $\mathbf{x}_{it}$ is a $p_x \times 1$ vector of independent variables, $\mathbf{f}_t$ is an $r \times 1$ vector of unobserved factors, γ_i is a vector of latent factor loadings, $\mathbf{\Gamma}_i$ is a $r \times p_x$ matrix of factor loadings, $\sigma(\cdot)$ is an unknown parameter associated with a conditional heteroscedastic model, u_{it} is an identically and independently distributed error term, and $\mathbf{v}_{it}$ is a p_x-dimensional vector assumed to follow an stationary process independently distributed of u_{it}. The variables α_i and μ_i are individual effects and α_i is an individual specific effect potentially correlated with the independent variables.

The above formulation include several panel data models investigated in the literature. The case of exogenous regressors with $\lambda_i = 0$ has been discussed in Pesaran (2006). In a series of recent papers, Chudik and Pesaran (2015) and Chudik, Mohaddes, Pesaran and Raissi (2015) investigate the estimation of a heterogeneous dynamic panel data model with strictly exogenous regressors and weakly exogenous covariates. Harding and Lamarche (2014) propose a quantile regression approach for the estimation of model (2.1) under the assumption that $\lambda_i = 0$ for all i. Moreover, Galvao (2011) proposes an instrumental variable approach for a dynamic quantile regression model when both $\gamma_i = \mathbf{0}$ and $\mathbf{\Gamma}_i = \mathbf{0}$. It is worth noticing that $\mathbf{f}_t$ is a vector of latent common factors and λ_i is assumed to satisfy $|\lambda_i| < 1$, and consequently, existing quantile regression approaches are inconsistent for the estimation of $(\lambda_i, \beta_i')'$ for $i = 1, \ldots, N$.

We estimate the following conditional quantile function:

$$Q_{Y_{it}}(\tau | y_{it-1}, \mathbf{x}_{it}, \alpha_i, \boldsymbol{f}_t) = \alpha_i(\tau) + \lambda_i(\tau) y_{it-1} + \mathbf{x}_{it}' \beta_i(\tau) + \mathbf{f}_t' \gamma_i(\tau), \tag{2.4}$$

where τ is a quantile in the interval $(0, 1)$ and the conditional quantile function is defined as $Q_{Y_{it}}(\tau | y_{it-1}, \mathbf{x}_{it}, \alpha_i, \mathbf{f}_t) := \inf\{y : P(Y_{it} < y | y_{it-1}, \mathbf{x}_{it}, \alpha_i, \mathbf{f}_t) \geq \tau\}$. The term $\mathbf{f}_t' \gamma_i(\tau)$ can be interpreted as a quantile-specific function capturing unobserved heterogeneity that was not adequately controlled by the independent variables in model (2.1). The model can be considered to be semi-parametric since the functional form of the conditional distribution of Y_{it} given $(y_{it-1}, \mathbf{x}_{it}', \alpha_i, \mathbf{f}_t')'$ is left unspecified and we do not impose a parametric assumption on the relation between the independent variables and the latent variables in the model. If we assume for instance that $\sigma(\mathbf{x}_{it}, \alpha_i, \mathbf{f}_t) = \mathbf{x}'_{it}\phi_x + \mathbf{f}'_t\phi_f$, the heteroscedastic form $\sigma(\mathbf{x}_{it}, \alpha_i, \mathbf{f}_t)u_{it}$ in equation (2.2) leads to a conditional quantile function with coefficients $\alpha_i(\tau) = \alpha_i$, $\lambda_i(\tau) = \lambda_i$, $\beta_i(\tau) = \beta_i + \phi_x Q_u(\tau)$, and $\gamma_i(\tau) = \gamma_i + \phi_f Q_u(\tau)$, where $Q_u(\tau) = F_u^{-1}(\tau)$ is the quantile function and F_u is the distribution of the error term u.

In this paper, we are interested in estimating the contemporaneous effect of a change in $\mathbf{x}_{it}$ on the quantiles of the conditional distribution of the response variable as well as its long run effect. For instance, in Section 4, we estimate an autoregressive panel quantile model for energy consumption with interactive effects. Our primary focus is to identify and estimate the effect of different technologies that enable households to respond to time-of-use pricing on energy consumption, focusing on the distributional effect of the assigned technologies. We also investigate the long-run effect of a change in energy price for different enabling technologies that can be defined as $\theta_i(\tau) = \beta_i(\tau)/(1 - \lambda_i(\tau)))$.

Naturally, the model can accommodate additional lags of the dependent variable, deterministic trends, time-invariant covariates, and lags of the exogenous covariates. These variations can be incorporated at a cost of additional notational complexity.

2.1. Estimation

We estimate the parameters of interest in equation (2.4) considering a quantile regression estimator for a dynamic model with interactive effects. The following assumptions are needed for the estimation of model (2.4).

Assumption 1. *$(u_{it}, \boldsymbol{v}'_{it})'$ satisfies $(u_{it}, \boldsymbol{v}'_{it})' = \sum_{l=0}^{\infty} \boldsymbol{S}_{il}\boldsymbol{\zeta}_{i,t-l}$, where $\boldsymbol{\zeta}_{it}$ is a vector of identically, independently distributed random variables with mean zero, variance matrix $\boldsymbol{I}_{k_1+1}$, and finite fourth order cumulants. In particular,*

$$\| Var(\mathbf{v}_{it})\| = \left\| \sum_l \boldsymbol{S}_{il}\boldsymbol{S}'_{il} \right\| \leq k < \infty.$$

Assumption 2. *The $r \times 1$ vector $\mathbf{f}_t = (f_{1t}, f_{2t}, \ldots, f_{rT})'$ is a zero mean, unit variance, and covariance stationary process, with absolute summable autocovariances, and it is distributed independently of $u_{it'}$ and $\mathbf{v}'_{it}$ for all i, t, t'.*

Assumption 3. *The factor loadings $\gamma_i = \gamma + \eta_{\gamma_i}$ and $vec(\boldsymbol{\Gamma}_i) = vec(\boldsymbol{\Gamma}) + \eta_{\Gamma_i}$ are distributed independently of u_{jt} and $\boldsymbol{v}_{jt}$ for all i and j with mean γ and $\boldsymbol{\Gamma}$ and bounded variances. The error* *terms η_{γ_i} and η_{Γ_i} are independent and identically distributed random variables with mean zero and covariances $\mathbf{\Omega}_\gamma$ and $\mathbf{\Omega}_\eta$, with $\|\mathbf{\Omega}_\gamma\| < K$ and $\|\mathbf{\Omega}_\eta\| < K$.*

Assumption 4. *The variables $\mathbf{x}_{it} = (x_{it,1}, x_{it,2}, \dots, x_{it,p})' \in \mathcal{X} \subset R^{p_x}$ and u_{it} are stochastically independent and generated according to equation (2.3).*

Assumption 5. *The support of λ_i for $i = 1, \dots, N$ lies (strictly) within the unit circle and the $p_x + 1$-dimensional vector of coefficients $\vartheta_i = (\lambda_i, \beta_i')'$ follows a random coefficient representation,*

$$\vartheta_i = \vartheta + \nu_i, \quad \nu_i \sim IID\left(\mathbf{0}, \mathbf{\Omega}_\vartheta\right), \tag{2.5}$$

for $i = 1, \dots, N$ where $\vartheta = (\lambda, \beta')'$, $\|\vartheta\| < K$, $\|\mathbf{\Omega}_\vartheta\| < K$ and $\mathbf{\Omega}_\vartheta$ is a symmetric positive definite matrix. The variables λ_i and β_i are independently distributed. Moreover, the random variable ν_i is independently distributed of γ_i, $\mathbf{\Gamma}_i$, u_{it}, $\mathbf{v}_{it}'$ and $\mathbf{f}_t$ for all i, t.

Assumption 6. *Let $\mathbf{C} = E(\mathbf{C}_i) = (\gamma, \mathbf{\Gamma})'$. Provided that $p_x > r - 1$, the $(p_x + 1) \times r$ dimensional matrix $\mathbf{C}$ has full column rank.*

In general, the assumptions are similar to the conditions in Pesaran (2006) and Chudik and Pesaran (2015). One difference is that it is common to assume that exists an N-dimensional vector of non-stochastic weights that satisfy granularity conditions. They are important in small samples but do not affect the asymptotic results established below in Section 2.2. Therefore, we implicitly consider the case of equal weights $1/N$. Moreover, it is worth mentioning that the full rank condition 6 ensures the large N representation of the unobserved factors, which is derived in what follows.

Let L be the lag operator. By Assumption 5, $(1 - \lambda_i L)$ is invertible for all $i = 1, \dots, N$, and then equation (2.1), after multiplying by $(1 - \lambda_i L)^{-1}$, can be written as,

$$y_{it} = \sum_{l=0}^{\infty} \lambda_i^l \alpha_i + \sum_{l=0}^{\infty} \lambda_i^l \beta_i' \mathbf{x}_{it-l} + \sum_{l=0}^{\infty} \lambda_i^l \gamma_i' \mathbf{f}_{t-l} + \sum_{l=0}^{\infty} \lambda_i^l \sigma(\cdot) u_{it-l}. \tag{2.6}$$

We now derive a large N representation for a linear combination of the latent factors following closely Pesaran (2006) and Chudik and Pesaran (2015). Taking cross-sectional averages in equation (2.6), we have that,

$$\bar{y}_t = \bar{\alpha} + a(L)\left(\beta' \bar{\mathbf{x}}_t + \gamma' \mathbf{f}_t\right) + O_p(N^{-1/2}), \tag{2.7}$$

where $\bar{w} = N^{-1} \sum_{i=1}^{N} w_{it}$ for a generic w_{it}, $\bar{\alpha} = N^{-1} \sum_{i=1}^{N} (1 - \lambda_i)^{-1} \alpha_i$, and $a(L) = \sum_{l=0}^{\infty} a_l L^l$, where $E(\lambda_i^l) = a_l$ for $l \in \{0, 1, 2, \dots\}$. Similarly, taking cross-sectional averages in equation (2.3), we obtain,

$$\bar{\mathbf{x}}_t = \bar{\mu} + \mathbf{\Gamma}' \mathbf{f}_t + O_p(N^{-1/2}), \tag{2.8}$$

where $\bar{\mu} = N^{-1}\sum_{i=1}^{N}\mu_i$. Letting $b(L) = a^{-1}(L)$ denote the inverse of $a(L)$, which exists by Assumption 5 and has exponentially decaying coefficients, and solving for $\mathbf{f}_t$ in equations (2.7) and (2.8), we write,

$$\mathbf{C}\mathbf{f}_t = \mathbf{\Lambda}(L)\bar{\mathbf{z}}_t - \bar{\mathbf{d}} + O_p(N^{-1/2}), \tag{2.9}$$

where $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{\mathbf{x}}_t')'$, $\mathbf{C} = E(\mathbf{C}_i) = (\gamma, \mathbf{\Gamma})'$, $\bar{\mathbf{d}} = (b(L)\bar{\alpha}, \bar{\mu}')'$ and,

$$\mathbf{\Lambda}(L) = \begin{bmatrix} b(L) & -\beta' \\ \mathbf{0} & \mathbf{I}_{p_x} \end{bmatrix}.$$

It follows that when the rank of $\mathbf{C}$ is equal to the number of factors r, we obtain the desired representation of the latent factors:

$$\mathbf{f}_t = \mathbf{f}_0 + \mathbf{G}(L)\bar{\mathbf{z}}_t + O_p(N^{-1/2}), \tag{2.10}$$

where $\mathbf{f}_0 = -(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\bar{\mathbf{d}}$ and $\mathbf{G}(L) = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{\Lambda}(L)$. Finally, substituting the representation of the factors in equation (2.1), we obtain that,

$$y_{it} = \beta_{0i} + \lambda_i y_{it-1} + \beta_i'\mathbf{x}_{it} + \gamma_i'\mathbf{G}(L)\bar{\mathbf{z}}_t + u_{it} + O_p(N^{-1/2}), \tag{2.11}$$

where $\beta_{0i} = \alpha_i - \gamma_i'\mathbf{f}_0$. Equation (2.11) leads to a conditional quantile function that is naturally different than equation (2.4) since $\mathbf{f}_t$ is unknown and we use a large N representation. Thus, the conditional quantile function that conditions on cross-sectional averages, $Q_{Y_{it}}(\tau|\cdot, \bar{\mathbf{z}}_t)$, approximates well the conditional quantile function for known factors, $Q_{Y_{it}}(\tau|\cdot, \mathbf{f}_t)$, when the number of cross-sectional units N is sufficiently large. To see this, assume $\sigma(\cdot) = \sigma = 1$ for simplicity and let $V_{it} := \gamma_i'\mathbf{H}\bar{\mathbf{V}}_t$ where $\mathbf{H} = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'$ and $\bar{\mathbf{V}}_t = (\bar{u}_t, \bar{\mathbf{v}}_t')'$. Noting that the error term in (2.11) can be written as $u_{it} + V_{it}$, we have that

$$\begin{aligned} u_{it} + V_{it} &= u_{it} + \gamma_i'\mathbf{H}\bar{\mathbf{V}}_t = u_{it} + \gamma_i'\left(\gamma\gamma' + \mathbf{\Gamma}\mathbf{\Gamma}'\right)^{-1}(\gamma\bar{u}_t + \mathbf{\Gamma}\bar{\mathbf{v}}_t) \\ &= u_{it} + \theta_{1i}\bar{u}_t + \theta_{2i}\bar{\mathbf{v}}_t, \end{aligned}$$

where $\theta_{1i} = \gamma_i'(\gamma\gamma' + \mathbf{\Gamma}\mathbf{\Gamma}')^{-1}\gamma$ and $\theta_{2i} = \gamma_i'(\gamma\gamma' + \mathbf{\Gamma}\mathbf{\Gamma}')^{-1}\mathbf{\Gamma}$. It follows that,

$$u_{it} + V_{it} = \left(\frac{N + \theta_{1i}}{N}\right)u_{it} + \frac{\theta_{1i}}{N}\sum_{j\neq i}^{N} u_{jt} + \frac{\theta_{2i}}{N}\sum_{i=1}^{N}\mathbf{v}_{it}. \tag{2.12}$$

Under Assumption 1 and Assumption 3 for each $1 \leq i \leq N$, the second and third term in equation (2.12) converge in probability to zero as $N \to \infty$ and thus the conditional quantile $Q_{u_{it}+V_{it}}(\tau) \approx$

$Q_{u_{it}}(\tau)$ for large N under Assumption 3. To avoid small sample biases from using the approximation of the latent factors in a quantile regression model, N needs to be sufficiently large.[1]

Remark 1. The simplest case when $p_x = 1 = r$ and $\beta = \mathbf{\Gamma} = \mathbf{0}$ offers insights on the observables approximating the latent factors in an autoregressive model. Assuming also that $\mathbf{f}_t$ and $\mathbf{x}_{it}$ are independent, and we have that,

$$\gamma_i \mathbf{G}(L)\bar{\mathbf{z}}_t = \frac{\gamma_i^2}{\gamma_i^2 + \Gamma_i^2} b(L)\bar{y}_t - \frac{\gamma_i^2 \beta + \gamma_i \Gamma_i}{\gamma_i^2 + \Gamma_i^2} \bar{x}_t = b(L)\bar{y}_t \tag{2.13}$$

suggesting that lagged values of the cross-sectional average of the dependent variable are important to approximate the latent factors $\mathbf{f}_t$.

We now present an approach that can be used to estimate a dynamic quantile regression model with interactive effects. First, based on equation (2.11), it is natural to consider the following cross-sectional augmented regression:

$$y_{it} = \lambda_i y_{it-1} + \beta_i' \mathbf{x}_{it} + \sum_{l=0}^{p_T} \delta_{il} \bar{y}_{t-l} + \delta_{ix}' \bar{\mathbf{x}}_t + e_{it}, \tag{2.14}$$

where the (p_x+1) dimensional vector δ_{ix} is a reduced form coefficient for the cross-sectional average of $\mathbf{x}_{it}$ and the error term e_{it} includes u_{it}, a term $O_p(N^{-1/2})$ associated with the approximation of $\mathbf{f}_t$ and an error component due to the possible infinite distributed lag function $\delta_i(L)$. Note that now $\bar{\mathbf{x}}_t$ is defined as a vector of independent variables that includes an intercept. Moreover, the number of lags is denoted by p_T and it is assumed that $p_{T_i} = p_T$ for all i for the simplicity of exposition. It is also assumed that the number of lags to approximate the factors is known and that $E(\lambda_i^l)$ decay exponentially which is satisfied by Assumption 5.

The quantile regression procedure is similar in spirit to Pesaran (2006), Harding and Lamarche (2014) and Chudik and Pesaran (2015). Define the parameter $\pi_i(\tau) := (\lambda_i(\tau), \beta_i'(\tau), \delta_{iy}'(\tau), \delta_{ix}'(\tau))'$ with $\delta_{iy}(\tau) = (\delta_{i1}(\tau), \delta_{i2}(\tau), \ldots, \delta_{ip_T}(\tau))'$, and

$$C_{it}(\tau, \pi_i) = \rho_\tau \left(y_{it} - \lambda_i y_{it-1} - \beta_i' \mathbf{x}_{it} - \sum_{l=0}^{p_T} \delta_{il} \bar{y}_{t-l} - \delta_{ix}' \bar{\mathbf{x}}_t \right), \tag{2.15}$$

[1] The Supplement offers finite sample evidence on the effect of conditioning on cross-sectional averages. It presents results from considering a series of relatively small number cross-sectional units.

where $\rho_\tau(u) = u(\tau - I(u \le 0))$ is the standard quantile regression loss function. First, we minimize the individual specific objective function (2.15) for $\pi_i(\tau)$,

$$\hat{\pi}_i(\tau) = \arg\min_{\pi_i \in \mathbf{\Pi}_i} \sum_{t=p_T+1}^{T} C_{it}(\tau, \pi_i). \tag{2.16}$$

Therefore, the quantile regression estimator for heterogeneous effects in a dynamic panel quantile with interactive effects, $\hat{\pi}_i(\tau)$, is based on the cross-sectionally augmented regression (2.14). We also propose a quantile group estimator for $\vartheta(\tau) := E((\lambda_i(\tau), \beta_i(\tau)')')$. The estimator is,

$$\hat{\vartheta}(\tau) = \frac{1}{N}\sum_{i=1}^{N} \hat{\vartheta}_i(\tau) = \frac{1}{N}\sum_{i=1}^{N} \left(\mathbf{\Xi}_i \circ \hat{\pi}_i(\tau)\right), \tag{2.17}$$

where $\circ$ denotes Hadamard product, $\mathbf{\Xi}_i = (\iota_i', \mathbf{0}_i')'$ with ι_i denoting a $p_x + 1$ dimensional vector of ones and $\mathbf{0}_i$ a $p_T + p_x + 1$ dimensional vector of zeros. In what follows, we denote the estimator defined in (2.17) as mean quantile group estimator, MQG.

2.2. Asymptotic Theory

This section investigates the large sample properties of the proposed estimators in equation (2.16) and equation (2.17). We consider the following regularity conditions for the consistency of the proposed estimators. Throughout this section, the vector of variables $\mathbf{X}_{it} = (y_{it-1}, \mathbf{x}_{it}', \bar{\mathbf{z}}_t')'$ and $\bar{\mathbf{z}}_t' = (\bar{y}_t, ..., \bar{y}_{t-p_T}, \bar{\mathbf{x}}_t')'$ and write equation (2.14) as $Y_{it} = \mathbf{X}_{it}'\pi_i + e_{it}$. Also $\|\cdot\|_1$ stands for the ℓ_1-norm.

Assumption 7. *Suppose that for each i, $\{(\boldsymbol{X}_{it}', Y_{it}) : t = 1, 2, \ldots\}$ is a stationary series and β-mixing time series with β-mixing coefficients $\beta_i(j)$. We assume that $\sup_{1 \le i \le N} \beta_i(j) \le Ba^j$ for all $j \ge 1$, $a \in (0,1)$ and $B > 0$. Moreover, $\{(\boldsymbol{X}_{it}', Y_{it}) : t = 1, 2, \ldots\}$ are independent across i.*

Assumption 8. *There exists a constant M such that $\max \|\boldsymbol{X}_{it}\| < M$.*

Assumption 9. *For each $\eta > 0$,*

$$\epsilon_\eta := \inf_{i \ge 1} \inf_{\|\boldsymbol{\theta}\|_1 = \eta} E\left[\int_0^{\boldsymbol{X}_{i1}'\boldsymbol{\pi}} \left(F_i(s|\boldsymbol{X}_{i1}) - \tau\right) ds\right],$$

where F_i is defined as a conditional distribution of u_{it} and the conditional densities f_i is continuous, uniformly bounded away from 0 and ∞, with continuous derivatives everywhere. Moreover, the joint distribution of (u_{i1}, u_{i1+j}), $f_{ij}(u_{i1}, u_{i1+j}|\mathbf{X}_{i1}, \mathbf{X}_{i1+j}) \le C_f$ with $C_f > 0$, uniformly over $(u_{i1}, u_{i1+j}, \mathbf{X}_{i1}, \mathbf{X}_{i1+j})$ for all $i \le 1$ and $j \le 1$.

These conditions are used in the literature. For instance, Assumption 7 has been used in Hahn and Kuersteiner (2004), Kato, Galvao and Montes-Rojas (2012), and Galvao, Lamarche and Lima (2013). The condition allows for dependence across time, implying that we need to apply a Bernstein type inequality for β-mixing sequences (Corollary C.1. in Kato, Galvao and Montes-Rojas (2012)) rather than a Hoeffding's inequality to show weak consistency. It also imposes stationarity. Assumption 8 is also common in the literature and is important for the finite dimensional uniform convergence of the objective function. This assumption can be relaxed using a moment condition as in Fernandez-Val (2005) and Kato, Galvao and Montes-Rojas (2012). For the consistency of the estimator, Kato et al. assume $\sup_{i\geq 1} E(\|\boldsymbol{x}_{i1}\|^{2s}) < \infty$ for some real $s \geq 1$ letting to a rate $N/T^s \to 0$ as $N \to \infty$. Assumption 8 is also needed for the limiting distribution of the estimator obtained in below in Theorem 3. Lastly, Assumption 9 is an identification condition and is similar to Assumptions (A3) and (B3) in Kato, Galvao and Montes-Rojas (2012). The assumption also imposes conditions on the joints distributions because the data can be non-independently distributed.

The following result states the weak consistency of the estimator:

Theorem 1 (Consistency of $\hat{\pi}_i(\tau)$)**.** *Suppose Y_{it} for $i = 1, ..., N$ and $t = 1, ..., T$ is given by the panel data model* (2.1)-(2.3) *and Assumptions 1-9 hold. As N, T and p_T go jointly to infinity with $p_T^3/T \to 0$ and $(\log(N))^2/T \to 0$, the quantile regression estimator for a model with interactive effects, $\hat{\pi}_i(\tau)$, is consistent.*

The proof is in Appendix A. As suspected, different conditions lead to changes in Theorem 1. Although we naturally need $N \to \infty$ for consistency of our approach, it is important to point out that if N is fixed, then $\sqrt{T}(\hat{\pi}_i(\tau) - \pi_i(\tau))$ converges in distribution to a mean zero random variable with covariance $\mathcal{V}$, under $T \to \infty$ and $p_T^3/T \to 0$. The form of the limiting covariance matrix depends on condition 7 and under i.i.d. conditions, the asymptotic covariance matrix is similar to the ones obtained in Koenker (2005). Under less general conditions in Assumption 7, (i.e., $\lambda_i = 0$ and $\mathbf{f}_t$ not serially dependent), an application of Hoeffding's inequality leads to a bound in Theorem 1 that is $O(\exp(-T)) = o(N^{-1})$ which is satisfied when $\log(N)/T \to 0$.

Although the consistency of individual coefficients might not imply the consistency of mean group estimators as discussed in Chudik and Pesaran (2015), our next result crucially follows when Assumption 6 holds.

Theorem 2 (Consistency of $\hat{\vartheta}(\tau)$)**.** *Under the conditions of Theorem 1, as (N, p_T, T) go jointly to infinity with $p_T^3/T \to 0$ and $(\log(N))^2/T \to 0$, the mean quantile group estimator for a model with interactive effects is weakly consistent, $\hat{\vartheta}(\tau) - \vartheta(\tau) \to 0$.*

We now turn our attention to the asymptotic distribution of the proposed estimator. We consider the following additional regularity conditions:

Assumption 10. *We assume that $\inf_i \lambda_{\min}(E(\mathbf{X}_{it}\mathbf{X}_{it}')) > 0$ where $\lambda_{\min(\cdot)}$ denotes the smallest eigenvalue and $E(\mathbf{X}_{it}\mathbf{X}_{it}')$ is a positive definite matrix.*

Assumption 11. *Let $\mathbf{J}_i := E[f(0|\mathbf{X}_{it})\mathbf{X}_{it}\mathbf{X}_{it}']$, $\mathbf{D}_i := E[\psi_\tau(Y_{it} < \mathbf{X}_{it}'\pi_i)\mathbf{X}_{it}\mathbf{X}_{it}']$, $\mathbf{\Gamma}_i := \mathbf{\Xi}_i\mathbf{\Xi}_i'$, $\mathbf{M}_i := E[\Phi_{it}\mathbf{X}_{it}\mathbf{X}_{it}']$ and $\Phi_{it} := I(Y_{it} \leq \mathbf{X}_{it}'\pi(\tau)) - I(Y_{it} \leq \mathbf{X}_{it}'\pi_i(\tau))$. The following conditions hold:*

(a) *The limiting matrices $\mathbf{J} := \lim_{N\to\infty} \mathbf{J}_N$ where $\mathbf{J}_N = N^{-1}\sum_{i=1}^N \mathbf{\Gamma}_i\mathbf{J}_i$ and $\mathbf{D} := \lim_{N\to\infty} \mathbf{D}_N$ where $\mathbf{D}_N = N^{-1}\sum_{i=1}^N \mathbf{\Gamma}_i\mathbf{D}_i$ exist and are nonsingular. Define $\mathbf{V}_\psi = \mathbf{J}^{-1}\mathbf{D}\mathbf{J}^{-1}$ as a $(p_x+1)\times(p_x+1)$ limiting positive definite matrix.*

(b) *The random variables Φ_{it} are independently distributed across i with mean zero and variance Var$(\mathbf{\Phi}) \leq K < \infty$. The limiting matrix exist and is nonsingular: $\mathbf{M} := \lim_{N\to\infty} \mathbf{M}_N$ where $\mathbf{M}_N = N^{-1}\sum_{i=1}^N \mathbf{\Gamma}_i\mathbf{M}_i$. Define $\mathbf{V}_v = \mathbf{J}^{-1}\mathbf{M}\mathbf{J}^{-1}$ as a $(p_x+1)\times(p_x+1)$ limiting positive definite matrix.*

(c) *The limiting matrix $\mathbf{V}_\psi + \mathbf{V}_v$ exist and is nonsingular.*

Assumption 10 is standard in the quantile regression literature and it is analogous to Assumption 7.b and 7.c in Chudik and Pesaran (2015). It guarantees that the inverse of $E[f(0|\mathbf{X}_{it})\mathbf{X}_{it}\mathbf{X}_{it}']$ exists, and jointly with Assumption 9, it implies that these inverses are uniformly bounded across i. Assumption 11 has three parts. The first part is standard in the panel quantile literature and it is needed for the existence of limiting forms of positive definite matrices and to invoke a Central Limit Theorem. The second part relates to slope heterogeneity in a quantile framework. Assumption 11.b allows a general form of slope heterogeneity while guaranteeing that the covariance matrix of the QMG estimator is well defined. It implies that slope heterogeneity can be more general than a location shift $v_i(\tau) = v_i$ for all i, imposing a symmetry condition around $\vartheta(\tau)$. The last assumption, Condition 11.c, guarantees that the asymptotic covariance matrix of the estimator exists and the minimum eigenvalue of the matrix is bounded away from zero.

The following theorem establishes the asymptotic distribution of the mean quantile group estimator.

Theorem 3 (Asymptotic Distribution of $\hat{\vartheta}(\tau)$)**.** *Suppose Y_{it} for $i = 1, ..., N$ and $t = 1, ..., T$ is given by the panel data model* (2.1)-(2.3) *and Assumptions 1-11 hold. As $(N, p_T, T) \to \infty$ with $p_T^3/T \to 0$ and $N^2(\log(N))^3/T \to 0$, the mean group quantile regression estimator for a model with interactive effects,*

$$\sqrt{NT}(\hat{\vartheta}(\tau) - \vartheta(\tau)) \rightsquigarrow \mathcal{N}(\mathbf{V}_\psi + \mathbf{V}_v).$$

It is worth mentioning that the restriction on T is that it grows almost exponentially in N and gives enough degrees of freedom for estimation of the model. The restriction on T helps to eliminate biases in the asymptotic distribution created by the remainder term in the Bahadur representation of the individual coefficients. It should be noted that for fixed N, the estimator $\delta_T(\tau) = \sqrt{T}(\hat{\vartheta}(\tau) - \vartheta(\tau))$ is asymptotically a Gaussian random variable. However, because the approximation of the factors requires $N \to \infty$ and we let N and T go jointly to infinity, T has to grow considerably faster to eliminate biases from incidental parameters and truncation of possibly infinite lag polynomials.

The asymptotic covariance matrix has a new component originated by slope heterogeneity. It is important to emphasize however that the asymptotic covariance matrix can be consistently estimated using existing estimators. For large N and T, we define $\hat{u}_{it}(\tau) := Y_{it} - \mathbf{X}_{it}'\hat{\pi}_i(\tau)$, h_N to be a sequence of bandwidths such that $h_N \to 0$ as $N \to \infty$, and $K_{h_N}(u) = h_N^{-1}K(u/h_N)$ be a Kernel estimator (see Koenker (2005)). Then we can use the following estimators to consistently estimate $\mathbf{V}_\psi$,

$$\hat{\mathbf{J}}_i := \frac{1}{T}\sum_{t=1}^{T} K(\hat{u}_{it}(\tau))\mathbf{X}_{it}\mathbf{X}_{it}', \quad \hat{\mathbf{D}}_i := \frac{1}{T}\sum_{t=1}^{T} \hat{\sigma}_\psi^2(q)\mathbf{X}_{it}\mathbf{X}_{it}',$$

where,

$$\hat{\sigma}_\psi^2(q) := \tau(1-\tau) + 2\sum_{j=1}^{q-1}\left(1 - \frac{j}{q}\right)\left[I(Y_{i1} \le \mathbf{X}_{i1}'\hat{\pi}_i(\tau), Y_{i1+j} \le \mathbf{X}_{i1+j}'\hat{\pi}_i(\tau)) - \tau^2\right].$$

To guarantee the consistent of the estimator, we require that $K(\cdot)$ is continuous and bounded, and $\log(N)/(Th_N) \to 0$ as $N \to \infty$. On the other hand, the asymptotic variance $\mathbf{V}_v$ can be estimated consistently considering,

$$\hat{\mathbf{V}}_v := \frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i\sum_{t=1}^{T}\hat{\mathbf{J}}_i^{-1}\mathbf{X}_{it}'\hat{\mathbf{\Phi}}_{it}\mathbf{X}_{it}\hat{\mathbf{J}}_i^{-1},$$

where $\hat{\mathbf{\Phi}}_{it} := (I(Y_{it} \le \mathbf{X}_{it}'\hat{\pi}(\tau)) - I(Y_{it} \le \mathbf{X}_{it}'\hat{\pi}_i(\tau)))$. Notice that if $\bar{v}_i(\tau, \mathbf{X}_i) = v_i$ for all i, the estimator $\hat{\mathbf{V}}_v$ is quantile invariant and therefore a consistent estimator of $\mathbf{V}_v$ can be defined as in equation (32) in Chudik and Pesaran (2015). In this case, the matrix $\hat{\mathbf{V}}_v$ can be estimated using $(N-1)^{-1}\sum_{i=1}^{N}(\hat{\vartheta}_i(\tau) - \hat{\vartheta}(\tau))(\hat{\vartheta}_i(\tau) - \hat{\vartheta}(\tau))'$.

3. Monte Carlo

This section reports results of several simulation exercises designed to evaluate the small sample performance of the proposed estimator. Observations on y_{it} for $i = 1, 2, \ldots, N$ and $t = -S + 1, -S + 2, .., 0, 1, ..., T$ are generated according to the following ARX(1) model with one factor:

$$y_{it} = \alpha_i + \lambda_i y_{i,t-1} + \beta_0 + \beta_i x_{it} + \gamma_i f_t + (1 + \delta x_{it})u_{it}, \tag{3.1}$$

where the error term u_{it} is distributed as $F(0, \sigma_i^2)$, for σ_i^2 generated as $0.5(1 + \chi_i)$ and χ_i denotes an i.i.d. random variable distributed as χ^2 distribution with 1 degree of freedom. Depending on the value of δ, we have two conditional quantile functions. When $\delta = 0$, we have that:

$$Q_{Y_{it}}(\tau|\alpha_i, y_{it-1}, x_{it}, f_t) = \alpha_i + \lambda_i y_{i,t-1} + \beta_0(\tau) + \beta_i x_{it} + \gamma_i f_t, \tag{3.2}$$

with $\beta_0(\tau) = \beta_0 + F_u^{-1}(\tau)$. On the hand, when $\delta \neq 0$, we have that:

$$Q_{Y_{it}}(\tau|\alpha_i, y_{it-1}, x_{it}, f_t) = \alpha_i + \lambda_i y_{i,t-1} + \beta_0(\tau) + \beta_i(\tau) x_{it} + \gamma_i f_t, \tag{3.3}$$

with $\beta_0(\tau) = \beta_0 + F_u^{-1}(\tau)$ and $\beta_i(\tau) = \beta_i + \delta F_u^{-1}(\tau)$. Models (3.2) and (3.3) are typically refereed as location shift and location-scale shift models in the literature (see, e.g., Koenker (2005)). Quantile regression models are estimated with an overall intercept β_0 which is assumed to be zero in the simulations. Note that for S sufficiently large, we have that

$$y_{i0} \approx \frac{1}{1-\lambda}\alpha_i + \beta \sum_{j=0}^{S-1} \lambda^j x_{i,-j} + \sum_{j=0}^{S-1} \lambda^j \xi_{i,-j}, \tag{3.4}$$

where $\xi_{it} = \gamma_i f_t + (1 + \delta x_{it})u_{it}$, $\lambda = E(\lambda_i)$ and $\beta = E(\beta_i)$. In all the variants of the model considered in the simulations, we assume that $S = 200$. The independent variable, x_{it}, is generated as

$$x_{it} = \mu_i + \Gamma_i f_t + z_{it}, \tag{3.5}$$

$$z_{it} = \rho_x z_{i,t-1} + \sqrt{1 - \rho_x^2}\varepsilon_{it}, \tag{3.6}$$

$$f_t = \rho_f f_{t-1} + \sqrt{1 - \rho_f^2}\varepsilon_{ft}, \tag{3.7}$$

where the i.i.d. variables $\mu_i \sim \mathcal{N}(0.5, 1)$, $\varepsilon_{it} \sim \mathcal{N}(0, 1)$, and $\varepsilon_{f_t} \sim \mathcal{N}(0, 1)$. We consider the case of relatively persistent regressors by assuming $\rho_x = 0.8$ and $\rho_f = 0.9$. Moreover, without loss of generality we assume $x_{i,-S} = 0$ and $f_{-S} = 0$.

The factor loadings in equation (3.1), γ_i, and in equation (3.5), Γ_i, are generated as $\gamma_i \sim iid\mathcal{N}(0.5, 1)$ and $\Gamma_i \sim iid\mathcal{N}(0.5, 1)$. Finally, the fixed effects, α_i, are allowed to be correlated with the errors by generating them as $\alpha_i = \bar{x}_i + \gamma_i \bar{f} + \bar{u}_i + v_i$, where the individual specific averages are defined as

$\bar{x}_i = T^{-1}\sum_{t=1}^{T} x_{it}$, $\bar{f} = T^{-1}\sum_{t=1}^{T} f_t$, $\bar{u}_i = T^{-1}\sum_{t=1}^{T} u_{it}$. The error term v_i in the equation for α_i is assumed to be distributed as $\mathcal{N}(0,1)$.

Without loss of generality, we set $\lambda_i = \lambda$ for all $i =, 1, \ldots, N$ and we consider three values of $\lambda = \{0.25, 0.50, 0.75\}$. Later in Figure 3.1, we investigate the performance of the estimator wwith heterogeneous λ_i's. Moreover, in addition to the experiments presented in this section, we also considered static panel data experiments (i.e., $\lambda = 0$) and compare the performance of the proposed approach with existing panel quantile regression approaches. For relatively large T, the performance of the proposed estimator was similar in both the static panel data model and dynamic panel data model. Thus, we present results for the dynamic model to save space.

We consider that the error term u_{it} is an i.i.d. random variable distributed as Gaussian, t-student with 4 degrees of freedom, and χ^2_3 with 3 degrees of freedom in the following 4 variations of the model:

Design 1: (Location shift model with homogeneous slopes). We assume $\beta = 1$ in a location shift model with $\delta = 0$.
Design 2: (Location shift model with heterogeneous slopes). We consider heterogeneous slope parameters $\beta_i = \beta + b_i$ in a location shift model, where $\delta = 0$, $\beta = 1$ and $b_i \sim \mathcal{U}(-0.25, 0.25)$. The parameter $\beta_i(\tau) = \beta_i$ for all i and τ.
Design 3: (Location-scale shift model with homogeneous slopes). We assume homogenous slope parameters $\beta = 1$ in a location-scale shift model with $\delta = 0.1$. In this case, the slope parameter $\beta(\tau) = \beta + 0.1F_u^{-1}(\tau)$.
Design 4: (Location-scale shift model with heterogeneous slopes). We consider heterogeneous slope parameters as in Design 2, $\beta_i = \beta + b_i$, in a location-scale shift model with $\delta = 0.1$. We assume $\beta = 1$ and $b_i \sim \mathcal{U}(-0.25, 0.25)$ which implies that $\beta_i(\tau) = 1 + b_i + 0.1F_u^{-1}(\tau)$. In this case, $E(\beta_i(\tau)) = \beta(\tau) = 1 + 0.1F_u^{-1}(\tau)$.

Tables 3.1 to Table 3.2 present the bias and root mean square error (RMSE) for the slope parameter β in the location shift model with $\lambda = 0.5$, which is similar to the values we find in Section 4. The finite sample performance for the slope parameter when the model 3.1 include a different value for λ is considered in the Supplement. While Table 3.1 presents results for Designs 1 and 2, Table 3.2 presents results for Designs 3 and 4. The tables show results for quantile regression estimators at two quantiles, $\tau \in \{0.25, 0.50\}$, based on sample sizes of $N \in \{100, 200\}$ and $T \in \{50, 100, 200\}$.

We compare the performance of the following quantile regression estimators: (i) the existing quantile regression estimator for a dynamic panel data model developed by Galvao (2011), labeled DQR, and (ii) the quantile mean group (QMG) estimator for a model with interactive effects. The DQR estimator uses y_{it-2} as an instrument for y_{it-1}. It should be noted that Galvao's model does not include the term $\lambda_i f_t$, which can generate biases that cannot be eliminated by the use of instrumental variables. The proposed quantile mean group estimator, QMG, is obtained as the cross sectional average of $\hat{\beta}_i(\tau)$ using $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{y}_{t-1}, \bar{x}_t)'$.

The tables do not provide the finite sample performance of other existing quantile estimators. The classical quantile regression estimator is biased because the individual specific effects α_i and the factor f_t are correlated with the independent variable x_{it}. Also the fixed effects estimator, a recently proposed minimum distance quantile regression estimator, and the penalized quantile regression estimator are biased when the model includes a lagged dependent variable. Therefore, we restrict attention to estimators for dynamic panel quantile regression models.

3.1. **Bias and Mean Square Error**

In Table 3.1, it might not be surprising to find out that the DQR method is biased and that its bias tends to be slightly larger in the case of heterogeneous slopes. The bias of this estimator for the slope β tends to increase as T increases and it does not seem to change at the 0.5 and 0.25 quantiles. On the other hand, the performance of the QMG estimator is excellent, with biases in general lower than 5% for $T = 50$ and biases decreasing rapidly to 1% when $T = 200$. In all the variations of the model considered in the table, the quantile estimator QMG performs better than DQR in terms of RMSE too.

Table 3.2 presents results for the location-scale shift model where $\beta(\tau)$ changes by quantile. For instance, $\beta(0.5) = 1$ and $\beta(0.25) = 0.93$ in the case where the error term $u_{it} \sim \mathcal{N}(0, 1)$, and $\beta(0.5) = 1.24$ and $\beta(0.25) = 1.12$ when $u_{it} \sim \chi^2_3$. We continue to see that the DQR estimator is biased and has poor MSE properties. The performance of the QMG estimator in these variations of the model is similar to Table 3.1, with low biases and small MSE. For values of T larger than 50, the bias of the proposed estimator is always negative and it ranges between 0.6% and 3%.

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
Design 1: Location shift with homogeneous slopes														
100	50	Bias	0.126	-0.052	0.125	-0.053	0.111	-0.050	0.111	-0.056	0.080	-0.066	0.080	-0.035
100	50	RMSE	0.145	0.054	0.143	0.055	0.127	0.051	0.128	0.057	0.102	0.068	0.098	0.037
100	100	Bias	0.163	-0.024	0.161	-0.024	0.150	-0.022	0.149	-0.025	0.113	-0.029	0.106	-0.014
100	100	RMSE	0.173	0.025	0.171	0.025	0.161	0.024	0.161	0.027	0.125	0.030	0.117	0.015
100	200	Bias	0.177	-0.008	0.177	-0.008	0.169	-0.007	0.169	-0.009	0.134	-0.013	0.123	-0.004
100	200	RMSE	0.182	0.010	0.182	0.010	0.174	0.009	0.174	0.011	0.141	0.014	0.129	0.006
200	50	Bias	0.127	-0.057	0.125	-0.056	0.120	-0.053	0.119	-0.060	0.082	-0.068	0.081	-0.037
200	50	RMSE	0.144	0.058	0.142	0.057	0.140	0.054	0.139	0.061	0.099	0.069	0.096	0.037
200	100	Bias	0.160	-0.026	0.159	-0.025	0.139	-0.023	0.139	-0.026	0.120	-0.032	0.111	-0.015
200	100	RMSE	0.171	0.026	0.169	0.026	0.148	0.024	0.148	0.027	0.130	0.032	0.120	0.016
200	200	Bias	0.187	-0.011	0.186	-0.011	0.166	-0.010	0.168	-0.011	0.137	-0.015	0.126	-0.006
200	200	RMSE	0.192	0.012	0.191	0.011	0.171	0.011	0.173	0.012	0.143	0.015	0.131	0.007
Design 2: Location shift with heterogeneous slopes														
100	50	Bias	0.131	-0.053	0.129	-0.053	0.112	-0.049	0.113	-0.055	0.083	-0.066	0.083	-0.035
100	50	RMSE	0.151	0.054	0.148	0.055	0.129	0.051	0.130	0.057	0.106	0.068	0.103	0.036
100	100	Bias	0.173	-0.023	0.170	-0.023	0.152	-0.022	0.151	-0.025	0.115	-0.029	0.109	-0.014
100	100	RMSE	0.181	0.025	0.180	0.024	0.163	0.023	0.163	0.026	0.127	0.030	0.119	0.015
100	200	Bias	0.185	-0.009	0.184	-0.009	0.174	-0.007	0.175	-0.009	0.138	-0.012	0.128	-0.004
100	200	RMSE	0.190	0.010	0.189	0.011	0.180	0.009	0.180	0.011	0.144	0.014	0.133	0.006
200	50	Bias	0.130	-0.057	0.128	-0.056	0.124	-0.054	0.122	-0.061	0.094	-0.068	0.092	-0.036
200	50	RMSE	0.147	0.058	0.145	0.057	0.143	0.054	0.142	0.061	0.110	0.069	0.106	0.037
200	100	Bias	0.160	-0.025	0.159	-0.025	0.144	-0.024	0.144	-0.027	0.120	-0.032	0.113	-0.015
200	100	RMSE	0.171	0.026	0.169	0.026	0.152	0.024	0.152	0.027	0.129	0.032	0.122	0.016
200	200	Bias	0.186	-0.011	0.185	-0.010	0.172	-0.010	0.174	-0.011	0.139	-0.015	0.128	-0.007
200	200	RMSE	0.190	0.012	0.189	0.011	0.178	0.011	0.179	0.012	0.144	0.015	0.132	0.007

TABLE 3.1. *Bias and root mean square error (RMSE) of quantile regression estimators for β in Designs 1 and 2. In all the variations of the model, $\lambda = 0.5$.*

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
Design 3: Location-scale shift with homogeneous slopes														
100	50	Bias	0.128	-0.059	0.125	-0.059	0.111	-0.056	0.110	-0.063	0.060	-0.062	0.060	-0.033
100	50	RMSE	0.146	0.061	0.143	0.061	0.131	0.058	0.130	0.065	0.080	0.064	0.078	0.035
100	100	Bias	0.170	-0.026	0.167	-0.026	0.143	-0.024	0.140	-0.027	0.093	-0.030	0.088	-0.013
100	100	RMSE	0.180	0.027	0.178	0.027	0.151	0.025	0.149	0.029	0.103	0.031	0.097	0.015
100	200	Bias	0.185	-0.010	0.183	-0.010	0.165	-0.009	0.164	-0.011	0.112	-0.013	0.102	-0.005
100	200	RMSE	0.190	0.012	0.188	0.012	0.171	0.010	0.171	0.013	0.117	0.014	0.107	0.007
200	50	Bias	0.120	-0.063	0.116	-0.063	0.113	-0.058	0.111	-0.065	0.061	-0.064	0.063	-0.033
200	50	RMSE	0.143	0.064	0.138	0.064	0.130	0.059	0.127	0.066	0.082	0.065	0.080	0.034
200	100	Bias	0.158	-0.029	0.156	-0.028	0.145	-0.026	0.143	-0.029	0.092	-0.029	0.085	-0.014
200	100	RMSE	0.168	0.030	0.166	0.029	0.154	0.027	0.152	0.030	0.101	0.030	0.093	0.014
200	200	Bias	0.185	-0.012	0.183	-0.012	0.163	-0.011	0.161	-0.013	0.109	-0.014	0.099	-0.006
200	200	RMSE	0.191	0.013	0.188	0.013	0.168	0.012	0.166	0.013	0.114	0.014	0.104	0.007
Design 4: Location-scale shift with heterogeneous slopes														
100	50	Bias	0.131	-0.062	0.127	-0.060	0.119	-0.057	0.117	-0.064	0.061	-0.065	0.062	-0.034
100	50	RMSE	0.148	0.063	0.145	0.062	0.140	0.059	0.139	0.066	0.080	0.067	0.077	0.036
100	100	Bias	0.158	-0.026	0.156	-0.026	0.153	-0.023	0.152	-0.028	0.090	-0.029	0.085	-0.013
100	100	RMSE	0.169	0.027	0.167	0.028	0.162	0.025	0.162	0.029	0.099	0.030	0.093	0.015
100	200	Bias	0.181	-0.010	0.178	-0.010	0.168	-0.009	0.167	-0.011	0.111	-0.012	0.101	-0.004
100	200	RMSE	0.185	0.012	0.182	0.012	0.174	0.010	0.173	0.012	0.116	0.014	0.106	0.006
200	50	Bias	0.133	-0.062	0.131	-0.063	0.123	-0.059	0.119	-0.066	0.064	-0.065	0.065	-0.034
200	50	RMSE	0.151	0.063	0.149	0.064	0.140	0.060	0.136	0.067	0.086	0.065	0.082	0.035
200	100	Bias	0.164	-0.029	0.161	-0.029	0.150	-0.026	0.147	-0.029	0.094	-0.030	0.087	-0.014
200	100	RMSE	0.172	0.029	0.169	0.029	0.159	0.026	0.156	0.029	0.103	0.030	0.096	0.015
200	200	Bias	0.191	-0.012	0.189	-0.013	0.170	-0.011	0.170	-0.013	0.110	-0.014	0.101	-0.006
200	200	RMSE	0.196	0.013	0.194	0.013	0.175	0.012	0.175	0.014	0.115	0.015	0.106	0.007

TABLE 3.2. *Bias and root mean square error (RMSE) of quantile regression estimators for β in Designs 3 and 4. In all the variations of the model, $\lambda = 0.5$.*

We now expand the simulation evidence for the slope parameter β to consider different values of λ. Table S.1 presents results for $\lambda \in \{0.25, 0.75\}$ considering the same designs as in Tables 3.1 and 3.2 and $N = 100$ and $T = 200$. We considered a moderate N and large T panel because our application in Section 4 employs a data set with 779 households and 8639 time-series observations. We see that the QMG estimator continues to perform better than the DQR estimator. We also find that the performance of the QMG estimator is invariant to the choice of λ, at least in the simulations considered thus far. We do investigate the performance of the QMG estimator when $\lambda \to 1$ below.

We now turn our attention to the estimator for λ and $\theta = \beta/(1-\lambda)$. The estimator for θ is defined as $\hat{\beta}/(1-\hat{\lambda})$ and it is obtained by plugging in the quantile estimates corresponding to λ and β. We employ this method for the DQR and QMG estimators.

Tables 3.3, 3.4, 3.5 and 3.6 show the bias and RMSE of the DQR and QMG estimators for the parameters of interest. These four tables show results for the four different designs we consider in this section. Each table presents, in columns, the performance of the estimators at $\tau \in \{0.25, 0.50\}$ and in rows the different samples sizes and distributions for the error term. The upper block present results when u_{it} is distributed as $\mathcal{N}(0, 1)$, the middle panel shows results when $u_{it} \sim t_4$ and the lower block presents results when $u_{it} \sim \chi^2_3$. While $\lambda = 0.5$ does not change in these tables, the parameter of interest θ does change in the tables. For instance, $\theta(0.5) = 2 = \theta(0.25)$ in the Gaussian case in Table 3.3, $\theta(0.5) = 2.48$ and $\theta(0.25) = 2.24$ when $u_{it} \sim \chi^2_3$ in Table 3.5.

As before, the results indicate that the bias of the DQR estimator can be large, in particular for the long run coefficient θ. The bias of the QMG estimator is small in the tables and it tends to zero as T increases, as expected. For $T = 50$, however, we see that the bias of the DQR estimator is smaller than the bias of the QMG estimator in the case of χ^2_3 for the parameter λ (i.e., columns (1) and (2)). We also find that the QMG estimator has smaller variance than the DQR estimator which might be expected since the QMG estimator does not employ instrumental variables. Even in the few cases where the bias of the DQR estimator is smaller than the bias of the QMG estimator, the QMG estimator offers the best performance in terms of MSE.

A comparison between the results for the long-run effect in the location shift model reveals that estimating heterogeneous effects is more demanding than estimating homogeneous effects, as expected. However, the QMG estimator offers nearly zero biases for large N and large T. The DQR estimator is biased and its performance is not satisfactory in terms of both bias and MSE. The location-shift case, presented in Tables 3.5 and 3.6, reveals similar findings. Overall, when $\lambda = 0.5$,

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
			Normal Distribution							
100	50	Bias	-0.110	0.051	0.411	-0.043	-0.107	0.053	0.412	-0.028
100	50	RMSE	0.164	0.058	0.461	0.069	0.158	0.062	0.463	0.068
100	100	Bias	-0.168	0.029	0.493	-0.017	-0.163	0.028	0.495	-0.013
100	100	RMSE	0.197	0.033	0.520	0.036	0.193	0.034	0.526	0.036
100	200	Bias	-0.196	0.009	0.503	-0.005	-0.195	0.009	0.502	-0.005
100	200	RMSE	0.213	0.017	0.521	0.022	0.212	0.016	0.521	0.024
200	50	Bias	-0.109	0.058	0.416	-0.044	-0.105	0.058	0.416	-0.030
200	50	RMSE	0.156	0.062	0.456	0.060	0.152	0.061	0.464	0.053
200	100	Bias	-0.167	0.030	0.476	-0.021	-0.164	0.030	0.474	-0.014
200	100	RMSE	0.198	0.032	0.499	0.029	0.194	0.033	0.498	0.028
200	200	Bias	-0.213	0.014	0.525	-0.008	-0.212	0.013	0.526	-0.005
200	200	RMSE	0.227	0.016	0.541	0.017	0.226	0.016	0.542	0.018
			t_4 distribution							
100	50	Bias	-0.089	0.049	0.362	-0.042	-0.086	0.055	0.371	-0.031
100	50	RMSE	0.139	0.058	0.401	0.075	0.140	0.067	0.416	0.073
100	100	Bias	-0.149	0.027	0.454	-0.017	-0.146	0.029	0.460	-0.018
100	100	RMSE	0.183	0.032	0.489	0.039	0.180	0.036	0.498	0.042
100	200	Bias	-0.179	0.010	0.489	-0.003	-0.178	0.012	0.496	-0.002
100	200	RMSE	0.198	0.015	0.510	0.023	0.197	0.019	0.517	0.027
200	50	Bias	-0.103	0.054	0.390	-0.046	-0.099	0.061	0.400	-0.036
200	50	RMSE	0.159	0.058	0.429	0.064	0.155	0.065	0.443	0.082
200	100	Bias	-0.132	0.028	0.419	-0.020	-0.131	0.029	0.422	-0.021
200	100	RMSE	0.161	0.031	0.438	0.032	0.159	0.033	0.442	0.037
200	200	Bias	-0.179	0.012	0.469	-0.008	-0.180	0.014	0.479	-0.007
200	200	RMSE	0.196	0.015	0.485	0.019	0.197	0.017	0.495	0.022
			χ^2_3 distribution							
100	50	Bias	-0.035	0.075	0.320	-0.035	-0.037	0.041	0.309	-0.015
100	50	RMSE	0.109	0.092	0.367	0.111	0.104	0.058	0.351	0.084
100	100	Bias	-0.086	0.037	0.378	-0.015	-0.079	0.017	0.351	-0.007
100	100	RMSE	0.126	0.049	0.419	0.059	0.115	0.027	0.388	0.040
100	200	Bias	-0.123	0.018	0.409	-0.003	-0.109	0.006	0.373	0.000
100	200	RMSE	0.147	0.030	0.431	0.044	0.131	0.016	0.393	0.027
200	50	Bias	-0.035	0.072	0.321	-0.050	-0.038	0.039	0.310	-0.026
200	50	RMSE	0.098	0.082	0.354	0.088	0.090	0.049	0.343	0.061
200	100	Bias	-0.093	0.038	0.400	-0.022	-0.083	0.019	0.368	-0.008
200	100	RMSE	0.127	0.044	0.423	0.049	0.114	0.026	0.388	0.034
200	200	Bias	-0.122	0.020	0.430	-0.007	-0.108	0.009	0.391	-0.002
200	200	RMSE	0.141	0.026	0.445	0.030	0.127	0.014	0.405	0.021

TABLE 3.3. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ in Design 1. In all the variations of the model, $\lambda = 0.5$.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
						Normal Distribution				
100	50	Bias	-0.112	0.065	0.442	-0.015	-0.107	0.063	0.439	-0.004
100	50	RMSE	0.170	0.071	0.491	0.057	0.166	0.070	0.487	0.066
100	100	Bias	-0.179	0.035	0.526	0.002	-0.176	0.034	0.519	0.007
100	100	RMSE	0.210	0.041	0.554	0.033	0.208	0.040	0.547	0.036
100	200	Bias	-0.220	-0.006	0.488	-0.036	-0.220	-0.006	0.480	-0.035
100	200	RMSE	0.235	0.014	0.510	0.042	0.235	0.015	0.502	0.042
200	50	Bias	-0.110	0.065	0.438	-0.030	-0.107	0.064	0.428	-0.018
200	50	RMSE	0.158	0.068	0.476	0.051	0.154	0.067	0.471	0.046
200	100	Bias	-0.157	0.039	0.502	0.001	-0.155	0.038	0.500	0.005
200	100	RMSE	0.188	0.041	0.525	0.022	0.186	0.041	0.525	0.027
200	200	Bias	-0.214	0.013	0.510	-0.007	-0.212	0.012	0.507	-0.006
200	200	RMSE	0.229	0.015	0.522	0.017	0.227	0.015	0.520	0.017
						t_4 distribution				
100	50	Bias	-0.059	0.083	0.436	0.027	-0.057	0.089	0.446	0.038
100	50	RMSE	0.123	0.089	0.468	0.068	0.123	0.096	0.484	0.077
100	100	Bias	-0.119	0.062	0.532	0.057	-0.116	0.065	0.541	0.055
100	100	RMSE	0.161	0.065	0.561	0.066	0.159	0.068	0.574	0.067
100	200	Bias	-0.178	0.017	0.511	0.013	-0.177	0.020	0.519	0.014
100	200	RMSE	0.197	0.021	0.530	0.026	0.196	0.024	0.540	0.030
200	50	Bias	-0.092	0.070	0.428	-0.012	-0.085	0.077	0.440	-0.024
200	50	RMSE	0.151	0.073	0.465	0.045	0.145	0.081	0.482	0.299
200	100	Bias	-0.118	0.048	0.472	0.023	-0.117	0.050	0.475	0.021
200	100	RMSE	0.148	0.050	0.490	0.034	0.147	0.052	0.493	0.037
200	200	Bias	-0.183	0.012	0.481	-0.008	-0.185	0.014	0.488	-0.007
200	200	RMSE	0.200	0.014	0.497	0.018	0.201	0.017	0.505	0.022
						χ^2_3 distribution				
100	50	Bias	-0.051	0.066	0.308	-0.051	-0.053	0.032	0.299	-0.034
100	50	RMSE	0.123	0.085	0.362	0.133	0.116	0.051	0.347	0.088
100	100	Bias	-0.113	0.013	0.335	-0.062	-0.106	-0.006	0.310	-0.054
100	100	RMSE	0.146	0.034	0.382	0.084	0.135	0.023	0.352	0.068
100	200	Bias	-0.155	-0.008	0.362	-0.053	-0.142	-0.020	0.327	-0.051
100	200	RMSE	0.175	0.025	0.389	0.069	0.160	0.025	0.350	0.058
200	50	Bias	-0.060	0.055	0.341	-0.078	-0.064	0.024	0.321	-0.054
200	50	RMSE	0.109	0.066	0.376	0.129	0.103	0.037	0.355	0.076
200	100	Bias	-0.111	0.020	0.365	-0.060	-0.105	0.001	0.337	-0.045
200	100	RMSE	0.143	0.031	0.385	0.074	0.134	0.015	0.358	0.053
200	200	Bias	-0.133	0.018	0.424	-0.012	-0.119	0.007	0.388	-0.007
200	200	RMSE	0.150	0.024	0.439	0.032	0.134	0.013	0.402	0.022

TABLE 3.4. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ in Design 2. In all the variations of the model, $\lambda = 0.5$.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
						Normal Distribution				
100	50	Bias	-0.110	0.058	0.420	-0.048	-0.091	0.066	0.407	-0.011
100	50	RMSE	0.161	0.066	0.467	0.076	0.145	0.073	0.449	0.067
100	100	Bias	-0.183	0.032	0.502	-0.015	-0.166	0.035	0.460	0.000
100	100	RMSE	0.214	0.037	0.532	0.036	0.198	0.041	0.487	0.039
100	200	Bias	-0.208	0.013	0.525	-0.005	-0.190	0.016	0.485	0.005
100	200	RMSE	0.226	0.018	0.546	0.022	0.209	0.021	0.507	0.024
200	50	Bias	-0.100	0.065	0.409	-0.050	-0.082	0.074	0.379	-0.012
200	50	RMSE	0.158	0.068	0.462	0.063	0.142	0.077	0.421	0.047
200	100	Bias	-0.163	0.033	0.471	-0.024	-0.146	0.038	0.440	-0.006
200	100	RMSE	0.191	0.036	0.495	0.038	0.176	0.040	0.462	0.027
200	200	Bias	-0.212	0.014	0.520	-0.010	-0.194	0.017	0.477	0.000
200	200	RMSE	0.228	0.017	0.537	0.019	0.210	0.020	0.492	0.016
						t_4 distribution				
100	50	Bias	-0.087	0.054	0.377	-0.056	-0.073	0.069	0.357	-0.005
100	50	RMSE	0.144	0.063	0.425	0.083	0.137	0.079	0.395	0.223
100	100	Bias	-0.129	0.029	0.452	-0.017	-0.114	0.040	0.418	0.002
100	100	RMSE	0.157	0.035	0.480	0.042	0.146	0.046	0.444	0.041
100	200	Bias	-0.173	0.012	0.481	-0.005	-0.158	0.019	0.446	0.008
100	200	RMSE	0.194	0.018	0.501	0.024	0.181	0.024	0.464	0.025
200	50	Bias	-0.087	0.058	0.387	-0.051	-0.069	0.074	0.372	-0.014
200	50	RMSE	0.130	0.063	0.426	0.068	0.115	0.079	0.407	0.080
200	100	Bias	-0.139	0.030	0.439	-0.025	-0.124	0.039	0.405	-0.008
200	100	RMSE	0.168	0.033	0.461	0.035	0.154	0.042	0.424	0.029
200	200	Bias	-0.173	0.014	0.462	-0.009	-0.157	0.019	0.424	0.000
200	200	RMSE	0.188	0.016	0.475	0.020	0.173	0.022	0.436	0.020
						χ^2_3 distribution				
100	50	Bias	-0.033	0.082	0.274	-0.055	-0.025	0.060	0.260	0.008
100	50	RMSE	0.119	0.100	0.321	0.126	0.106	0.072	0.294	0.080
100	100	Bias	-0.091	0.044	0.354	-0.026	-0.067	0.025	0.326	0.003
100	100	RMSE	0.133	0.055	0.385	0.069	0.110	0.036	0.352	0.050
100	200	Bias	-0.136	0.017	0.367	-0.015	-0.102	0.010	0.320	0.002
100	200	RMSE	0.159	0.028	0.392	0.046	0.127	0.018	0.343	0.028
200	50	Bias	-0.037	0.082	0.275	-0.063	-0.025	0.057	0.274	0.001
200	50	RMSE	0.114	0.089	0.318	0.090	0.096	0.064	0.304	0.052
200	100	Bias	-0.091	0.041	0.343	-0.032	-0.064	0.024	0.312	-0.002
200	100	RMSE	0.127	0.047	0.364	0.054	0.100	0.029	0.329	0.033
200	200	Bias	-0.130	0.020	0.361	-0.015	-0.095	0.012	0.321	0.002
200	200	RMSE	0.152	0.026	0.377	0.035	0.117	0.017	0.335	0.022

TABLE 3.5. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ in Design 3. In all the variations of the model, $\lambda = 0.5$.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
						Normal Distribution				
100	50	Bias	-0.106	0.068	0.453	-0.038	-0.086	0.077	0.433	0.003
100	50	RMSE	0.156	0.072	0.504	0.064	0.142	0.082	0.480	0.061
100	100	Bias	-0.164	0.027	0.469	-0.021	-0.146	0.034	0.442	-0.003
100	100	RMSE	0.195	0.033	0.498	0.041	0.179	0.039	0.468	0.036
100	200	Bias	-0.201	0.011	0.514	-0.006	-0.181	0.015	0.477	0.005
100	200	RMSE	0.216	0.017	0.532	0.024	0.198	0.021	0.494	0.024
200	50	Bias	-0.116	0.069	0.443	-0.036	-0.097	0.078	0.428	-0.004
200	50	RMSE	0.169	0.072	0.484	0.055	0.152	0.081	0.464	0.044
200	100	Bias	-0.174	0.026	0.476	-0.035	-0.155	0.030	0.446	-0.019
200	100	RMSE	0.195	0.029	0.499	0.042	0.177	0.032	0.466	0.032
200	200	Bias	-0.215	0.014	0.554	-0.010	-0.196	0.018	0.516	0.001
200	200	RMSE	0.230	0.017	0.567	0.019	0.212	0.021	0.529	0.016
						t_4 distribution				
100	50	Bias	-0.092	0.066	0.411	-0.034	-0.076	0.086	0.393	0.007
100	50	RMSE	0.160	0.073	0.457	0.071	0.149	0.094	0.433	0.076
100	100	Bias	-0.153	0.021	0.462	-0.029	-0.138	0.033	0.432	-0.012
100	100	RMSE	0.182	0.029	0.494	0.046	0.169	0.040	0.462	0.042
100	200	Bias	-0.173	0.019	0.505	0.010	-0.155	0.026	0.470	0.023
100	200	RMSE	0.195	0.023	0.525	0.025	0.180	0.030	0.488	0.032
200	50	Bias	-0.109	0.050	0.390	-0.068	-0.090	0.069	0.364	-0.031
200	50	RMSE	0.153	0.055	0.428	0.080	0.138	0.073	0.397	0.058
200	100	Bias	-0.145	0.033	0.457	-0.015	-0.128	0.040	0.424	-0.001
200	100	RMSE	0.172	0.036	0.480	0.030	0.156	0.043	0.445	0.031
200	200	Bias	-0.182	0.015	0.488	-0.008	-0.167	0.019	0.455	0.000
200	200	RMSE	0.199	0.018	0.502	0.019	0.185	0.022	0.469	0.018
						χ^2_3 distribution				
100	50	Bias	-0.031	0.092	0.288	-0.057	-0.021	0.064	0.281	0.010
100	50	RMSE	0.108	0.107	0.344	0.111	0.097	0.076	0.320	0.073
100	100	Bias	-0.060	0.074	0.410	0.035	-0.034	0.056	0.385	0.064
100	100	RMSE	0.112	0.082	0.438	0.072	0.093	0.062	0.409	0.078
100	200	Bias	-0.106	0.051	0.445	0.052	-0.070	0.042	0.402	0.068
100	200	RMSE	0.134	0.056	0.467	0.067	0.102	0.045	0.422	0.074
200	50	Bias	-0.044	0.083	0.284	-0.064	-0.032	0.055	0.272	-0.004
200	50	RMSE	0.121	0.092	0.323	0.098	0.105	0.063	0.299	0.059
200	100	Bias	-0.101	0.041	0.334	-0.034	-0.073	0.023	0.308	-0.005
200	100	RMSE	0.137	0.047	0.356	0.055	0.108	0.028	0.325	0.031
200	200	Bias	-0.136	0.021	0.358	-0.014	-0.102	0.013	0.320	0.003
200	200	RMSE	0.157	0.026	0.374	0.031	0.123	0.017	0.334	0.020

TABLE 3.6. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ in Design 4. In all the variations of the model, $\lambda = 0.5$.*

the QMG estimator offers the best performance in terms of bias and MSE in the class of estimators for a dynamic quantile panel data model.

Figure 3.1 offers a clear summary of the small sample performance of the QMG estimator as λ increases. The figure shows the bias and RMSE of the QMG estimator at $\tau \in \{0.25, 0.50\}$ for λ, β and θ in terms of λ. We considered Design 1 with $N = 100$ and $T = 200$. Recall that when λ increases, θ increases too. For instance, while $\lambda = 0$ gives $\theta = \beta = 1$, $\lambda = 0.9$ gives $\theta = 10$ in our simulation experiment. Consistent with our previous evidence, we see that the performance of QMG estimator does not depend on λ when the interest is in estimating the covariate effect or short term effect, β. The bias tends to slightly increase but it is never larger than 1% for large values of λ. We also find that the RMSE of the estimator of β does not change with λ. On the other hand, we find that the absolute value of the bias of the QMG estimator for θ increases exponentially when $\lambda \to 1$. The figure shows that the bias, in absolute value, is negligible for $\lambda < 0.75$, and it increases rapidly when $\lambda > 0.8$. Note however that the bias in relative terms is always less than 10%. We also find that the RMSE increases with λ and that the RMSE of the QMG estimator at $\tau = 0.25$ is larger than the QMG estimator at $\tau = 0.50$, as expected.

Figure 3.1 also shows the bias and RMSE of the QMG estimator when $\lambda_i = \lambda + \omega_i$, where $\omega_i \sim \mathcal{U}[-0.025, 0.025]$ and λ takes values in the interval $\lambda \in [0.05, 0.90]$. The parametrization guarantees that θ exists for all values of λ_i for $i = 1, \ldots, N$. We generate data using Design 1 with $N = 100$ and $T = 200$. Consistent with our expectations, the bias and RMSE of the estimator tends to be similar to the case of homogeneous λ's, although the performance of the estimator appears to deteriorate for large values of λ. We see an increase in the variance of the estimator, but the bias for θ remains, in absolute value, small for $E(\lambda_i) < 0.65$. As it can be seen in Figure 3.1, the parameter $(E(\lambda_i), \beta)$ can be estimated with small bias and excellent MSE performance in the case of heterogeneous λ_i's.

3.2. **Inference**

We now turn our attention to the standard error of the QMG estimator for λ and β. Tables 3.7 and 3.8 report the average estimated standard error obtained by the procedure outlined in Section 2.2. We select $q = 3$ to minimize potential biases in the estimation of the standard errors. Table 3.7 reports the standard error for Designs 1 and 2, and Table 3.8 reports results for Designs 3 and 4. We also report the standard deviation of the estimator based on 400 Monte Carlo repetitions.

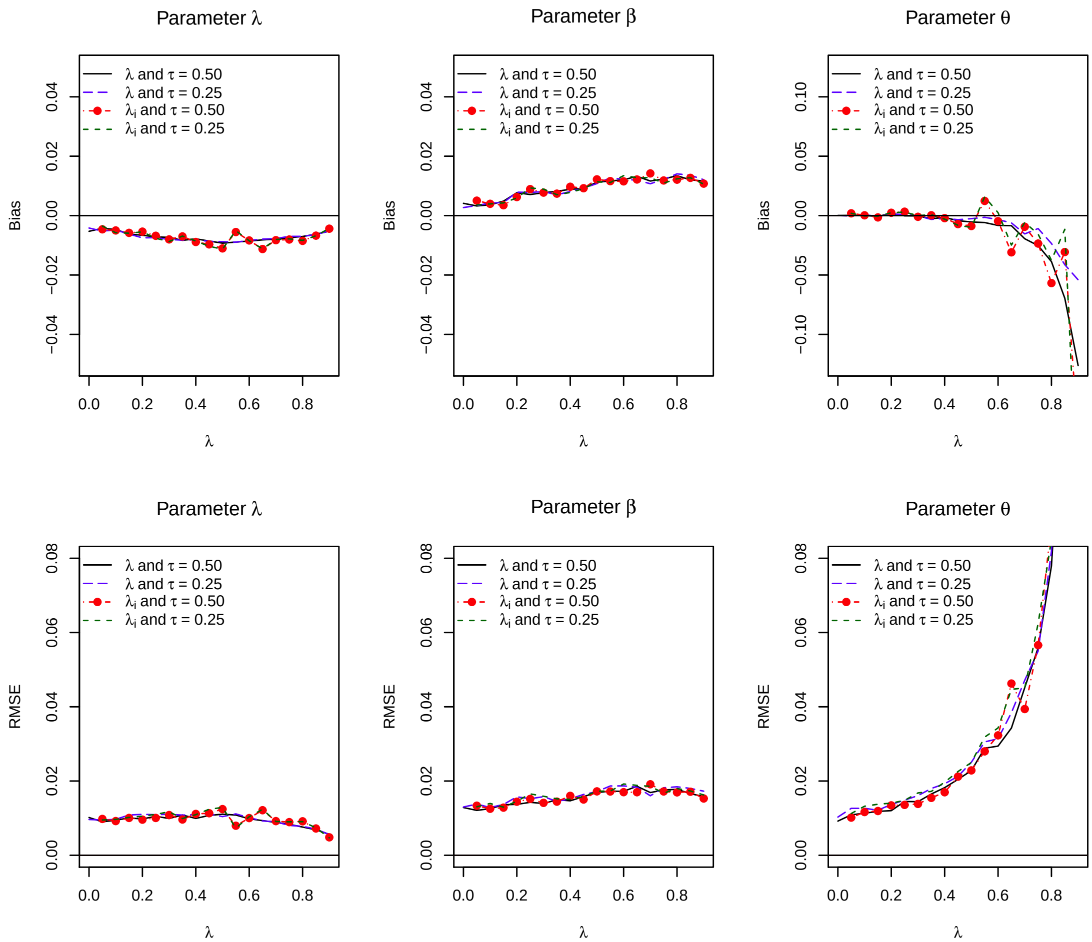

FIGURE 3.1. *Small sample performance of the QMG estimator for different values of* λ.

Because T relative to N is important for inference, we included results with $N \in \{50, 100\}$ and $T \in \{50, 100, 200\}$.

The results show that the estimated standard errors approximate very closely to the standard deviation of the estimator when T is larger than N. This result is expected by the rates of convergence needed to establish the consistency of the QMG estimator. The approximation is excellent in the case of the Normal and t_4 distributions. The evidence when $u_{it} \sim \chi^2_3$ suggest that the standard error needs a larger T relative to N to approximate well the truth.

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			λ		β		λ		β		λ		β	
			0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
Design 1: Location shift with homogeneous slopes														
50	50	Std error	0.020	0.021	0.053	0.053	0.020	0.022	0.060	0.063	0.024	0.020	0.099	0.079
50	50	Std dev	0.019	0.020	0.039	0.042	0.019	0.021	0.046	0.053	0.021	0.016	0.073	0.054
50	100	Std error	0.013	0.014	0.035	0.034	0.013	0.014	0.038	0.039	0.016	0.012	0.062	0.048
50	100	Std dev	0.013	0.014	0.026	0.029	0.013	0.014	0.029	0.032	0.013	0.010	0.046	0.032
50	200	Std error	0.009	0.010	0.024	0.024	0.009	0.010	0.025	0.027	0.011	0.008	0.041	0.031
50	200	Std dev	0.009	0.010	0.019	0.020	0.008	0.009	0.020	0.023	0.009	0.007	0.033	0.024
100	50	Std error	0.014	0.015	0.037	0.037	0.014	0.015	0.042	0.043	0.017	0.014	0.070	0.055
100	50	Std dev	0.013	0.014	0.027	0.029	0.012	0.014	0.029	0.035	0.014	0.011	0.051	0.039
100	100	Std error	0.009	0.010	0.024	0.024	0.009	0.010	0.026	0.027	0.011	0.009	0.044	0.033
100	100	Std dev	0.009	0.010	0.017	0.020	0.008	0.009	0.019	0.022	0.009	0.007	0.034	0.023
100	200	Std error	0.006	0.007	0.016	0.016	0.006	0.007	0.017	0.018	0.008	0.006	0.029	0.021
100	200	Std dev	0.006	0.006	0.013	0.013	0.005	0.007	0.013	0.015	0.006	0.005	0.022	0.016
Design 2: Location shift with heterogeneous slopes														
50	50	Std error	0.020	0.021	0.057	0.057	0.020	0.022	0.063	0.066	0.024	0.020	0.102	0.082
50	50	Std dev	0.019	0.021	0.043	0.049	0.018	0.020	0.047	0.054	0.020	0.015	0.074	0.059
50	100	Std error	0.013	0.014	0.040	0.040	0.013	0.014	0.043	0.044	0.016	0.012	0.066	0.052
50	100	Std dev	0.013	0.015	0.033	0.036	0.012	0.014	0.035	0.039	0.014	0.010	0.053	0.040
50	200	Std error	0.009	0.010	0.031	0.031	0.009	0.010	0.033	0.033	0.011	0.008	0.046	0.038
50	200	Std dev	0.010	0.010	0.029	0.029	0.009	0.010	0.029	0.031	0.010	0.007	0.038	0.033
100	50	Std error	0.014	0.015	0.040	0.040	0.014	0.016	0.044	0.046	0.017	0.014	0.072	0.057
100	50	Std dev	0.013	0.015	0.032	0.034	0.013	0.015	0.032	0.039	0.014	0.010	0.053	0.039
100	100	Std error	0.009	0.010	0.028	0.028	0.009	0.010	0.030	0.031	0.011	0.009	0.046	0.036
100	100	Std dev	0.008	0.010	0.023	0.025	0.008	0.009	0.024	0.028	0.009	0.007	0.037	0.029
100	200	Std error	0.006	0.007	0.021	0.021	0.006	0.007	0.022	0.023	0.008	0.006	0.032	0.026
100	200	Std dev	0.006	0.007	0.020	0.020	0.005	0.006	0.020	0.022	0.006	0.004	0.028	0.023

TABLE 3.7. *Standard error (Std error) of the QMG estimator for λ and β in Designs 1 and 2. Std dev denotes standard deviation and $\lambda = 0.5$ in the simulations.*

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			λ		β		λ		β		λ		β	
			0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
			Design 3: Location-scale shift with homogeneous slopes											
50	50	Std error	0.022	0.023	0.058	0.057	0.021	0.023	0.065	0.066	0.026	0.021	0.114	0.092
50	50	Std dev	0.021	0.021	0.042	0.043	0.019	0.022	0.045	0.048	0.019	0.016	0.078	0.059
50	100	Std error	0.014	0.015	0.038	0.037	0.014	0.015	0.041	0.042	0.017	0.013	0.072	0.056
50	100	Std dev	0.012	0.014	0.027	0.029	0.012	0.014	0.028	0.031	0.013	0.010	0.050	0.038
50	200	Std error	0.010	0.010	0.025	0.025	0.009	0.010	0.027	0.028	0.012	0.009	0.047	0.036
50	200	Std dev	0.010	0.010	0.020	0.020	0.008	0.010	0.020	0.025	0.009	0.007	0.034	0.025
100	50	Std error	0.015	0.016	0.040	0.040	0.015	0.016	0.045	0.047	0.018	0.015	0.080	0.064
100	50	Std dev	0.015	0.015	0.031	0.031	0.013	0.015	0.031	0.037	0.014	0.010	0.057	0.042
100	100	Std error	0.010	0.010	0.026	0.025	0.010	0.010	0.028	0.029	0.012	0.009	0.051	0.039
100	100	Std dev	0.009	0.010	0.019	0.020	0.008	0.009	0.020	0.024	0.009	0.006	0.035	0.025
100	200	Std error	0.007	0.007	0.017	0.017	0.006	0.007	0.018	0.019	0.008	0.006	0.033	0.025
100	200	Std dev	0.006	0.007	0.012	0.014	0.005	0.007	0.013	0.016	0.006	0.004	0.022	0.016
			Design 4: Location-scale shift with heterogeneous slopes											
50	50	Std error	0.022	0.023	0.061	0.060	0.021	0.023	0.068	0.070	0.026	0.021	0.115	0.094
50	50	Std dev	0.021	0.023	0.048	0.050	0.019	0.021	0.047	0.059	0.019	0.015	0.082	0.064
50	100	Std error	0.014	0.015	0.042	0.041	0.014	0.015	0.046	0.047	0.017	0.013	0.074	0.059
50	100	Std dev	0.014	0.015	0.036	0.037	0.012	0.014	0.037	0.040	0.012	0.009	0.054	0.042
50	200	Std error	0.010	0.010	0.032	0.032	0.009	0.010	0.034	0.034	0.012	0.009	0.051	0.041
50	200	Std dev	0.010	0.011	0.029	0.029	0.008	0.009	0.029	0.030	0.009	0.007	0.038	0.032
100	50	Std error	0.015	0.016	0.043	0.043	0.015	0.016	0.047	0.049	0.018	0.015	0.082	0.066
100	50	Std dev	0.015	0.015	0.030	0.034	0.013	0.015	0.036	0.041	0.014	0.010	0.056	0.043
100	100	Std error	0.010	0.010	0.030	0.029	0.010	0.010	0.032	0.032	0.012	0.009	0.052	0.041
100	100	Std dev	0.009	0.010	0.023	0.024	0.008	0.009	0.024	0.027	0.009	0.006	0.039	0.029
100	200	Std error	0.007	0.007	0.022	0.022	0.006	0.007	0.023	0.024	0.008	0.006	0.036	0.029
100	200	Std dev	0.006	0.007	0.019	0.020	0.006	0.007	0.019	0.022	0.006	0.004	0.027	0.021

TABLE 3.8. *Standard error (Std error) of the QMG estimator for λ and β in Designs 3 and 4. Std dev denotes standard deviation and $\lambda = 0.5$ in the simulations.*

		Normal Distribution				t_4 distribution				χ^2_3 distribution			
		λ		β		λ		β		λ		β	
		0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
Design 1: Location shift with homogeneous slopes													
50	50	0.320	0.350	0.913	0.893	0.320	0.295	0.935	0.905	0.200	0.628	0.965	0.998
50	100	0.768	0.723	0.963	0.950	0.720	0.655	0.983	0.955	0.665	0.903	0.973	0.985
50	200	0.913	0.905	0.985	0.988	0.950	0.923	0.988	0.980	0.915	0.980	0.983	0.993
100	50	0.028	0.040	0.783	0.735	0.010	0.023	0.840	0.793	0.020	0.223	0.900	0.968
100	100	0.265	0.355	0.865	0.813	0.318	0.300	0.930	0.888	0.243	0.695	0.918	0.978
100	200	0.755	0.795	0.948	0.948	0.798	0.718	0.978	0.948	0.623	0.938	0.960	0.983
Design 2: Location shift with heterogeneous slopes													
50	50	0.308	0.373	0.915	0.888	0.315	0.288	0.948	0.918	0.220	0.620	0.963	0.983
50	100	0.750	0.733	0.958	0.953	0.753	0.670	0.963	0.928	0.643	0.908	0.980	0.985
50	200	0.903	0.923	0.938	0.968	0.938	0.930	0.970	0.958	0.885	0.973	0.980	0.978
100	50	0.023	0.053	0.770	0.775	0.040	0.038	0.833	0.785	0.003	0.228	0.915	0.973
100	100	0.300	0.368	0.885	0.865	0.325	0.268	0.948	0.870	0.208	0.718	0.940	0.978
100	200	0.745	0.803	0.938	0.935	0.805	0.755	0.963	0.940	0.583	0.928	0.948	0.965
Design 3: Location-scale shift with homogeneous slopes													
50	50	0.270	0.315	0.905	0.863	0.263	0.278	0.950	0.878	0.313	0.740	0.978	0.998
50	100	0.675	0.693	0.970	0.903	0.683	0.660	0.983	0.925	0.723	0.938	0.983	0.993
50	200	0.928	0.920	0.988	0.978	0.940	0.905	0.990	0.948	0.935	0.985	0.995	0.995
100	50	0.020	0.035	0.750	0.628	0.013	0.010	0.860	0.670	0.030	0.340	0.885	0.943
100	100	0.215	0.260	0.853	0.763	0.258	0.213	0.910	0.740	0.260	0.798	0.953	0.978
100	200	0.675	0.688	0.943	0.873	0.768	0.658	0.978	0.900	0.703	0.938	0.990	0.995
Design 4: Location-scale shift with heterogeneous slopes													
50	50	0.245	0.308	0.895	0.848	0.280	0.268	0.965	0.878	0.293	0.725	0.955	0.975
50	100	0.708	0.688	0.955	0.905	0.703	0.615	0.973	0.920	0.770	0.940	0.983	0.990
50	200	0.908	0.905	0.955	0.938	0.948	0.915	0.980	0.958	0.913	0.978	0.993	0.985
100	50	0.025	0.025	0.795	0.678	0.023	0.025	0.865	0.740	0.035	0.320	0.925	0.950
100	100	0.245	0.278	0.895	0.788	0.273	0.243	0.923	0.823	0.343	0.838	0.950	0.988
100	200	0.705	0.745	0.958	0.933	0.703	0.653	0.960	0.905	0.720	0.953	0.970	0.985

TABLE 3.9. *Empirical coverage probability for a nominal 95 percent confidence interval.*

		Normal Distribution				t_4 distribution				χ^2_3 distribution			
		λ		β		λ		β		λ		β	
		0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25	0.50	0.25
Design 1: Location shift with homogeneous slopes													
50	50	0.998	0.995	0.085	0.105	0.998	0.998	0.085	0.083	1.000	0.993	0.010	0.063
50	100	0.998	0.998	0.683	0.650	0.995	0.995	0.580	0.515	0.998	0.998	0.130	0.388
50	200	1.000	1.000	0.993	0.995	0.998	0.998	0.980	0.970	1.000	1.000	0.613	0.945
100	50	1.000	1.000	0.175	0.203	1.000	1.000	0.115	0.098	1.000	1.000	0.015	0.128
100	100	1.000	1.000	0.925	0.908	1.000	1.000	0.883	0.825	1.000	1.000	0.230	0.760
100	200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.883	1.000
Design 2: Location shift with heterogeneous slopes													
50	50	0.995	0.995	0.080	0.103	0.998	1.000	0.073	0.063	1.000	1.000	0.008	0.060
50	100	1.000	0.998	0.503	0.503	0.998	1.000	0.448	0.360	0.998	0.995	0.128	0.360
50	200	0.998	0.998	0.885	0.885	1.000	1.000	0.863	0.835	1.000	1.000	0.510	0.785
100	50	1.000	1.000	0.175	0.195	1.000	1.000	0.115	0.113	1.000	1.000	0.015	0.100
100	100	1.000	1.000	0.790	0.778	1.000	1.000	0.755	0.673	1.000	1.000	0.268	0.675
100	200	1.000	1.000	0.993	0.995	1.000	1.000	0.990	0.968	1.000	1.000	0.783	0.973
Design 3: Location-scale shift with homogeneous slopes													
50	50	0.998	0.998	0.058	0.040	0.998	1.000	0.045	0.020	0.998	0.993	0.008	0.008
50	100	1.000	0.998	0.530	0.465	1.000	1.000	0.473	0.303	0.998	0.998	0.058	0.205
50	200	0.998	0.998	0.985	0.973	1.000	1.000	0.980	0.908	1.000	1.000	0.460	0.810
100	50	1.000	1.000	0.135	0.078	1.000	1.000	0.080	0.040	1.000	1.000	0.000	0.028
100	100	1.000	1.000	0.820	0.768	1.000	1.000	0.815	0.543	1.000	1.000	0.125	0.480
100	200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.820	1.000
Design 4: Location-scale shift with heterogeneous slopes													
50	50	0.998	1.000	0.058	0.050	1.000	0.998	0.053	0.040	1.000	0.998	0.005	0.030
50	100	0.998	0.995	0.503	0.420	0.998	0.998	0.390	0.285	1.000	1.000	0.073	0.175
50	200	0.995	0.998	0.855	0.808	0.998	1.000	0.808	0.710	1.000	1.000	0.375	0.650
100	50	1.000	1.000	0.083	0.065	1.000	1.000	0.098	0.053	1.000	1.000	0.005	0.020
100	100	1.000	1.000	0.715	0.608	1.000	1.000	0.698	0.480	1.000	1.000	0.170	0.453
100	200	1.000	1.000	0.985	0.985	1.000	1.000	0.988	0.948	1.000	1.000	0.713	0.950

TABLE 3.10. *Power performance of the QMG estimator. We compute the power for the estimation of λ and β with the alternatives $H_a : \lambda = 0.55$ and $H_a : \beta = 1.1$.*

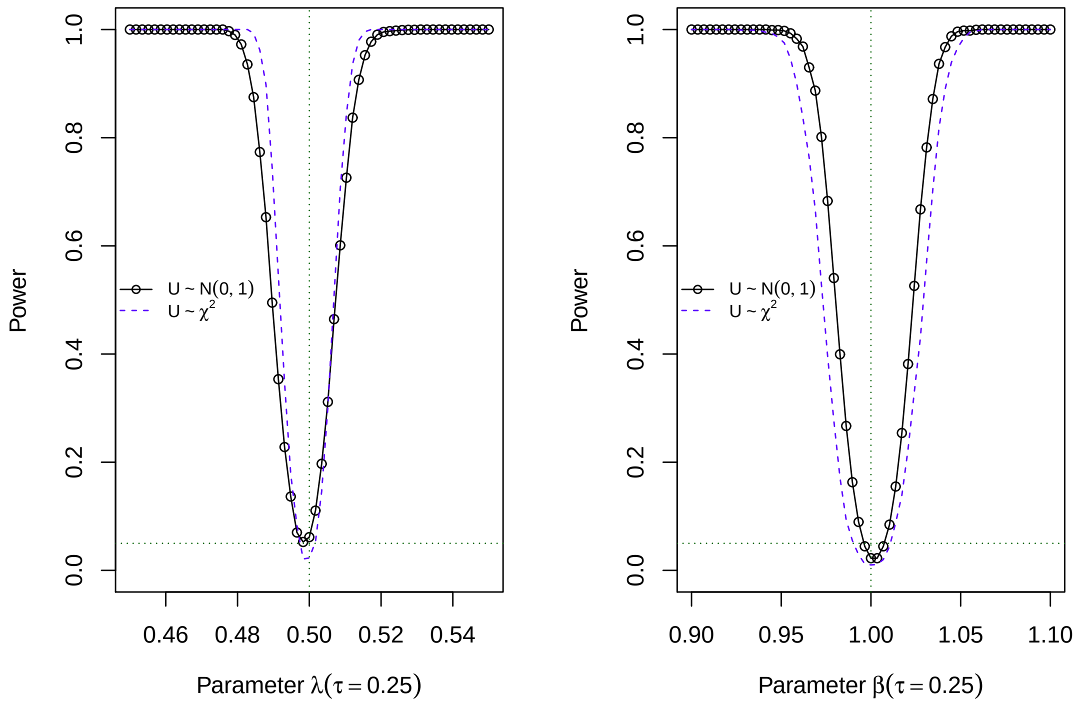

FIGURE 3.2. *Power of the QMG estimator against different alternatives.*

Table 3.9 provides empirical coverage probabilities for a nominal 95% confidence interval. The probabilities are calculated based on asymptotic Gaussian confidence intervals consistent with Theorem 3. We see different finite sample performances of the estimator for λ and β. If we examine the results across the different distributions, the QMG estimator does not perform well for λ when $T/N < 4$. On the other hand, the coverage probabilities for β approximate closely 0.95 with the exception when $T = 50$ and $N = 100$.

Lastly, we investigate the performance of the QMG estimator in terms of power. The results are shown in Table 3.10. We compute the power for the estimation of λ with the alternative hypothesis $H_a : \lambda = 0.55$ and β with the alternative hypothesis $H_a : \beta = 1.1$. The condition on the rate of convergence plays an important role in ensuring that the estimator has good power. In particular, the power is high for values of $T > 100$, although how quickly approximates 1 depends on the distribution of the error term and the number of cross-sectional units, N.

As we can see in Figure 3.2, the power function of the test constructed based on the QMG estimator at the 0.25 quantile tends to be symmetric and have the expected shape. As an illustration, Figure 3.2 reports results based on Design 1 for the Gaussian and χ^2_3 case when $N = 100$ and $T = 400$. Consistent with the theory, a larger T for a given number of cross-sectional units leads to a better approximation of the function (as documented in the Supplement). The evidence shows that for this sample size and quantile, the QMG estimator seem to perform reasonably well for different distributions and parameters.

4. Time-of-Use Pricing, Smart Technology and Energy Savings

Household responsiveness to Time-of-Use (TOU) energy pricing has been a topic of fundamental interest since the energy crisis during the 1970s. The recent academic literature has found substantial peak load reductions resulting from TOU pricing (Allcott (2011), Jessoe and Rapson (2014) and Ito (2014)). The literature also documents how consumer responsiveness to TOU prices is impacted by different types of enabling technologies (Jessoe and Rapson (2014)). Harding and Lamarche (2016) find that different technologies including smart meters and smart thermostats are associated with energy savings. They estimate the impact of TOU pricing using a randomized controlled trial of over 11 million observations on 15-minute interval electricity consumption in the US. In this section, we consider data from a similar randomized controlled trial, to study effectiveness of three major enabling technologies (web portal, in-home display and smart thermostat) within the context of TOU pricing.

We apply our quantile regression approach to estimate an autoregressive panel quantile model for energy consumption with interactive effects. We then compare the effect of different technologies on energy consumption, focusing on the distributional effect of these randomly assigned technologies. We find that smart thermostats are particularly effective relative to other enabling technologies at enabling households to respond to TOU pricing. The differential effects are more pronounced at the lower tail of the conditional distribution of energy consumption. This finding is interpreted as suggesting that smart thermostats, in addition of providing real time information on consumption and pricing, allow households to respond to price changes in advance by programming temperature settings for different times of the day. While households appear to reduce overall consumption as a result of these technologies relative to the control group, the average response does not summarize the distributional effect of the technologies. Lastly, we investigate the long-run effect of a change in energy price for different enabling technologies and we concentrate on the effects for different age and income groups.

4.1. **Data**

We employ data from a large scale randomized controlled trial (RCT) of TOU pricing for residential electricity consumers in a South Central US state. The data used in this paper includes 779 customers who were randomly assigned to a time-of-use pricing structure and received three different enabling technologies. The random allocation of a large sample of households into three treatment groups and one control group, and the availability of electricity readings measured over 15-minute intervals make the application of our QMG estimator particularly well suited to answer questions about the distributional effect of enabling technologies.

The experiment was conducted during the period of June 1st and September 30th of 2011. After households signed up for the program, they were randomized assigned into different treatment groups and a control group. Consumers randomly assigned to the control group were informed they were not eligible for the program. These households were kept on standard residential tariff and did not receive any new technology. On the other hand, customers who were selected to participate in the program were assigned to a time-of-use pricing rate which varied over two daily time periods. During the off-peak part of the day consisting of all hours except 2pm to 7pm, the rate charged for electricity consumption was \$0.042 kWh. During the on-peak part of the day, which was the period from 2pm to 7pm, the rate charged was \$0.23 kWh. Weekends were considered to be off-peak throughout.

We restrict the sample to households who did not change treatment status and whose electricity readings measured over 15-minute were consistently recorded in the period between June and September. This leads to a balanced panel of 6,729,781 observations with $T = 8,639$. The households in the treated groups had access to a website which exhibited information on their electricity consumption and prices. Our sample includes a group of 189 households who were limited to the website as the only enabling technology. As shown in Table 4.1, we have 1,632,771 observations for this group based on a large number of time series observations. The other households in the treatment group were randomly assigned to receive one of these two additional enabling technologies: in-home display (IHD) or "smart" programmable communicating thermostat (PCT). An IHD is a small wireless device which displays information on electricity usage and cost in real time. The PCT

	Control		Portal		IHD		PCT	
	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
Kilowatt-hours	0.61	0.51	0.62	0.52	0.59	0.48	0.59	0.48
Treatment	0.14	0.35	0.14	0.35	0.14	0.35	0.14	0.35
High Income	0.38	0.49	0.58	0.49	0.51	0.50	0.49	0.50
Low Income	0.31	0.46	0.21	0.40	0.18	0.39	0.23	0.42
Medium Income	0.31	0.46	0.22	0.41	0.31	0.46	0.28	0.45
Family Life	0.49	0.50	0.42	0.49	0.45	0.50	0.37	0.48
Mature	0.20	0.40	0.26	0.44	0.28	0.45	0.31	0.46
Young years	0.31	0.46	0.32	0.47	0.27	0.44	0.33	0.47
Temperature	84.88	12.85	84.85	12.85	84.89	12.85	84.95	12.85
Dew Point	58.51	7.91	58.53	7.93	58.50	7.91	58.43	7.88
Number of households	242		189		152		196	
Number of periods	8639		8639		8639		8639	
Number of observations	2090638		1632771		1313128		1693244	

TABLE 4.1. *Descriptive Statistics for the Smart Meter Data*

provides an interface that allows the customer to program and control the air conditioning system and respond to future and current price events. It also offers price and consumption information as the IHD screen. While a group of 152 households received in-home displays, another group of 196 customers received "smart" programmable communicating thermostats.

We have a limited number of observed covariates, mainly explained by confidentiality concerns which limit our access to household level information. We are able however to partition our sample by life stages and income. In Table 4.1, "young years" is designed to capture younger households, under 45 years of age with no children. The "family life" segment captures middle aged families with children. Households were also clustered by income into three groups: low, middle and high. The high group includes households with income above \$75,000 and the middle income group captures households with income between \$30,000 and \$75,000. As in Harding and Lamarche (2016), we use zip-code specific temperature and humidity records from the nearest weather station and demographic information provided by Nielsen's PRIZM® segmentation system.

4.2. Model

Recall that each household was randomly assigned to either a treatment group or the control group. Let denote groups by g and $g \in \{0, 1, 2, 3\}$. Hence, $g = 0$ denote the control group and $g \in \{1, 2, 3\}$ denote households assigned to either Portal, IHD or PCT. Let denote households by

$i = 1, \ldots, N_g$ and 15-minute intervals by $t = 1, \ldots, T$. We emphasize that t indicates a given household's electricity consumption continuously measured over 96 intervals per day over roughly 90 days. To explore the importance of heterogeneity of treatment effects, we estimate the following dynamic panel data model:

$$\log(Y_{it}) = \alpha_{ig} + \lambda_{ig}\log(Y_{it-1}) + \delta_{ig}d_t(g) + \mathbf{x}_{it}'\beta_{ig} + \gamma_{ig}'\mathbf{f}_t + u_{it}, \quad (4.1)$$

where Y_{it} is electricity usage and the vector $\mathbf{x}_{it}$ includes weather measurements. The variable $d_t(g)$ indicates the treatment assignment g and it takes the value 1 if t is between 2 pm and 7 pm during weekdays, and 0 otherwise. Our quantile treatment coefficients are identified by the time variation associated with TOU pricing:

$$Q_{\log(Y_{it})}(\tau|\cdot) = \alpha_{ig}(\tau) + \lambda_{ig}(\tau)\log(Y_{it-1}) + \delta_{ig}(\tau)d_t(g) + \mathbf{x}_{it}'\beta_{ig}(\tau) + \gamma_{ig}(\tau)'\mathbf{f}_t, \quad (4.2)$$

where $Q_{\log(Y_{it})}(\tau|\cdot)$ is the τ-th conditional quantile function and $\delta_{ig}(\tau)$ is the quantile treatment effect (QTE) of interest. We estimate the model using our QMG estimator for each quantile τ and group g separately. The standard errors are estimated using $\hat{\sigma}_{\psi}^2(3)$ as in Section 3.2 and the asymptotic covariance matrix obtained in Theorem 3.

4.3. Main Empirical Results

Table 4.2 reports results for the coefficient $\lambda_g(\tau) = E(\lambda_{ig}(\tau))$ and the QTE, $\delta_g(\tau) = E(\delta_{ig}(\tau))$, for the four groups: control group, portal, in-home display (IHD), and programmable communicating thermostats (PCT). The first two columns present results obtained by using the classical fixed effects (FE) estimator which produces inconsistent results in a model with interactive effects and the mean group (MG) estimator as in Chudik and Pesaran (2015). The next five columns show the quantile regression version of the MG estimator, labeled QMG.

The FE results tend to overestimate the effect of the lagged dependent variable and the treatment effect. Because these results are likely to be biased, we concentrate our attention on the MG estimator. The positive and significant coefficient for the control group indicates that consumption increases by 10% from 2 pm to 7 pm when temperature is likely to be high.[2] However, TOU pricing scheme seem to reduce energy consumption since the other treatment effects are smaller than 0.095. The table shows however that the technology adopted by households crucially determines whether the households engage in some saving behavior. The coefficient for Portal and IHD are positive

[2]The mean maximum daily temperature was 99°F and the median was 103°F. The months of July and August were very similar and September was substantially cooler with mean temperatures of 88.6°F.

	FE	MG	Quantiles				
			0.10	0.25	0.50	0.75	0.90
Control Group							
Consumption at	0.623	0.476	0.466	0.576	0.620	0.479	0.353
$t-1$ (in logs)	(0.001)	(0.009)	(0.020)	(0.021)	(0.021)	(0.020)	(0.015)
Treatment	0.145	0.095	0.148	0.112	0.063	0.049	0.048
(2pm - 7pm)	(0.001)	(0.018)	(0.011)	(0.009)	(0.006)	(0.007)	(0.006)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	242	242	242	242	242	242	242
$N \times T$	2090638	2090638	2090638	2090638	2090638	2090638	2090638
Portal							
Consumption at	0.622	0.489	0.469	0.588	0.631	0.485	0.360
$t-1$ (in logs)	(0.001)	(0.009)	(0.021)	(0.023)	(0.024)	(0.022)	(0.015)
Treatment	0.102	0.048	0.089	0.065	0.038	0.023	0.007
(2pm - 7pm)	(0.001)	(0.017)	(0.013)	(0.012)	(0.010)	(0.011)	(0.014)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	189	189	189	189	189	189	189
$N \times T$	1632771	1632771	1632771	1632771	1632771	1632771	1632771
IHD							
Consumption at	0.627	0.480	0.470	0.581	0.615	0.475	0.352
$t-1$ (in logs)	(0.001)	(0.009)	(0.022)	(0.025)	(0.027)	(0.025)	(0.018)
Treatment	0.089	0.046	0.102	0.072	0.038	0.025	0.006
(2pm - 7pm)	(0.002)	(0.017)	(0.017)	(0.013)	(0.010)	(0.009)	(0.011)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	152	152	152	152	152	152	152
$N \times T$	1313128	1313128	1313128	1313128	1313128	1313128	1313128
PCT							
Consumption at	0.771	0.682	0.718	0.785	0.806	0.693	0.560
$t-1$ (in logs)	(0.000)	(0.007)	(0.024)	(0.020)	(0.019)	(0.022)	(0.021)
Treatment	-0.010	-0.073	-0.092	-0.054	-0.029	-0.030	-0.029
(2pm - 7pm)	(0.001)	(0.014)	(0.022)	(0.016)	(0.010)	(0.010)	(0.014)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	196	196	196	196	196	196	196
$N \times T$	1693244	1693244	1693244	1693244	1693244	1693244	1693244

TABLE 4.2. *Quantile Mean Group estimator results for the control group and different technologies. FE denotes fixed effects and MG denotes the standard Mean Group estimator. IHD denotes in-home display and PCT is programmable communicating thermostats. Standard errors are in parentheses.*

and significant, and they suggest a decrease of 50% in energy use relative to the control group (although the differences might not be statistically significantly different than zero). However, the effect for the households using PCT are negative and significant relative to the other groups. The table shows that smart thermostats are particularly effective in enabling consumers to respond to TOU pricing. Households provided with a PCT achive a reduction of 7% when energy prices are high.

Households response however is not homogeneous across the quantiles of the conditional distribution of electricity consumption. Among consumers with a PCT technology, we find the largest energy saving in the lower tail of the conditional distribution, while the effect of TOU pricing is insignificant at the upper conditional quantile. In fact, the results seem to suggest that the three technologies are equally effective at the 0.9 quantile. When we examine the distributional effect among households with Portal and IHD technologies, we find a similar pattern. The QTE decreases in absolute value as we go across quantiles, changing from a significant effect at the 0.1 quantile to an effect not significantly different than zero at the 0.9 quantile. The effect of using PCT continues to be negative at the lower tail, and the effect of IHD is positive, although smaller than the estimate for the control group. This is an interesting finding that has policy implications as it suggests that consumers reacted to the change in price but an IHD is not as effective as the PCT in terms of energy savings.

Figure 4.1 offers a clear view of the main findings. The figure presents estimates of the QTE as a function of the quantile τ of the conditional distribution of electricity consumption. We present estimates for the short run and long run effects for the three treatment groups and the control group. The continuous line show the QMG estimates of $\delta_g(\tau)$ and the dashed line show QMG estimates for $\theta_g(\tau) = \delta_g(\tau)/(1 - \lambda_g(\tau))$ which is estimated as discussed above. The gray areas denote 95% point-wise confidence intervals.

We find that the estimates for the control group are displaced down when we compare then to the estimated corresponding to the treatment groups. This can be interpreted as suggesting that Portal and IHD reduce consumption during when electricity was more expensive but these technologies do not seem to achieve a significant energy reduction at any conditional quantile. On the other hand, the profiles of QTE for the control group and PCT group are clearly different, suggesting that smart thermostats are effective in allowing households to respond to the increase in the price of electricity between 2 pm and 7 pm. Moreover, it is interesting to see that the largest savings differentials in the short-run and long-run are estimated at the lower tail of the conditional distribution, while these

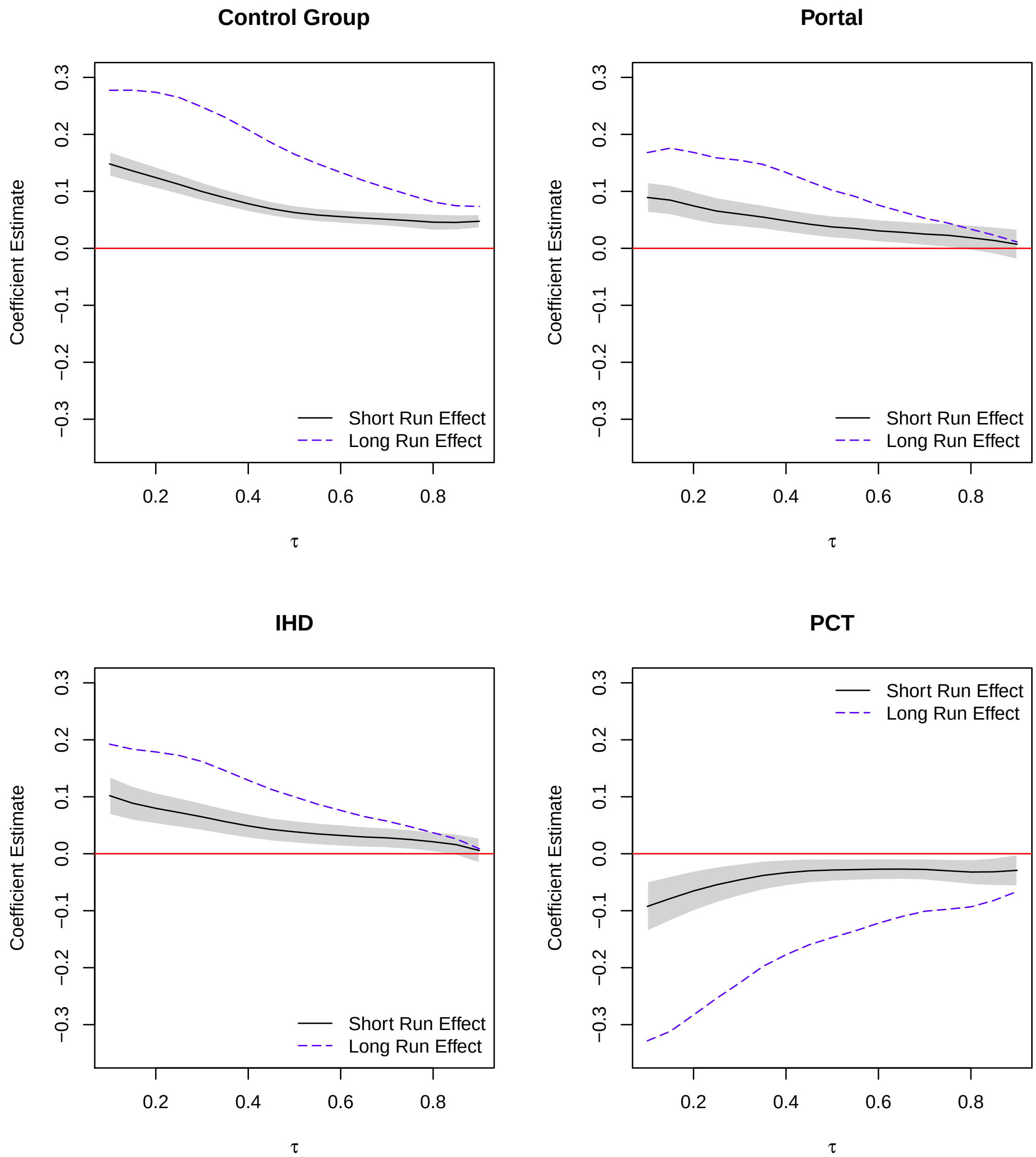

Figure 4.1. *Short and Long Run Quantile Regression Results*

differentials are small at the 0.9 quantile. The evidence indicates that households provided with a PCT can engage in considerable energy savings in the long run and the impact of the enabling technology is greater at the 0.1 quantile of the distribution of electricity consumption.

4.4. **Responsiveness across Demographics**

It is often important for policymakers to understand how the responsiveness to TOU pricing and enabling technologies changes with household demographics. This section address this question offering evidence on how consumers with different characteristics respond to TOU pricing. The household characteristics are limited to age and income of the family.

We first turn our attention to estimating the QTE across different levels of income. Table 4.3 is similar to Table 4.2 although it shows separate results for high- and low-income families. As discussed previously in Section 4.1, the high income group includes households with income above $75,000 and we combine the low and middle income groups to form a group of households with income below $75,000. As expected, high income households in the control group consume more electricity between 2 pm and 7 pm than low income households in the control group. The differential is fairly constant across quantiles. It is very interesting to discover that the results for the other groups are exactly the opposite: the coefficient estimates for high income consumers are smaller than the coefficient estimates for low income consumers. This suggests that high income customers are more successful in taking advantage of the existing information about price and consumption, and consequently, engage in larger electricity savings. This is true for all quantiles and groups. When we compare the evidence in Table 4.3 with the evidence presented in Table 4.2, we find that the effect of PCT continues to be negative but it is now significant at the 0.9 quantile for high income households. Thus, high-income customers who are conditionally consuming high levels of electricity reduce consumption by 4.3% relative to other times of the day and by roughly 10.3% relative to the control group in the period 2 pm to 7 pm. The energy savings among low-income households using PCT at the 0.9 quantile is nearly half of the percentage reduction of wealthier households.

Lastly, we investigate how households at different life stages respond to TOU pricing and the different technologies. In Table 4.4, the group called "family life" includes middle aged families with children, while "other years" refers to younger households under 45 years of age and no children and customers typically over 65 years of age. Again, as in the previous table, we see considerable response heterogeneity by group demographics. For instance, we find larger energy

		Quantiles					MG
		0.10	0.25	0.50	0.75	0.90	
				Control Group			
High Income	Consumption at	0.470	0.583	0.624	0.492	0.368	0.484
	$t-1$ (in logs)	(0.032)	(0.035)	(0.036)	(0.034)	(0.025)	(0.009)
	Treatment	0.164	0.132	0.075	0.061	0.058	0.110
	(2pm - 7pm)	(0.017)	(0.015)	(0.010)	(0.009)	(0.009)	(0.016)
Low Income	Consumption at	0.463	0.572	0.617	0.472	0.344	0.472
	$t-1$ (in logs)	(0.025)	(0.026)	(0.026)	(0.024)	(0.018)	(0.009)
	Treatment	0.139	0.100	0.055	0.041	0.041	0.085
	(2pm - 7pm)	(0.014)	(0.011)	(0.008)	(0.009)	(0.009)	(0.019)
				Portal			
High Income	Consumption at	0.463	0.582	0.619	0.480	0.351	0.483
	$t-1$ (in logs)	(0.029)	(0.031)	(0.033)	(0.030)	(0.020)	(0.009)
	Treatment	0.083	0.061	0.035	0.016	-0.003	0.040
	(2pm - 7pm)	(0.018)	(0.018)	(0.015)	(0.017)	(0.022)	(0.017)
Low Income	Consumption at	0.476	0.597	0.648	0.492	0.371	0.496
	$t-1$ (in logs)	(0.030)	(0.033)	(0.035)	(0.032)	(0.024)	(0.009)
	Treatment	0.098	0.072	0.042	0.032	0.021	0.060
	(2pm - 7pm)	(0.020)	(0.016)	(0.012)	(0.009)	(0.012)	(0.018)
				IHD			
High Income	Consumption at	0.470	0.582	0.609	0.481	0.364	0.488
	$t-1$ (in logs)	(0.032)	(0.036)	(0.038)	(0.037)	(0.025)	(0.009)
	Treatment	0.092	0.062	0.025	0.011	-0.015	0.028
	(2pm - 7pm)	(0.023)	(0.021)	(0.017)	(0.015)	(0.020)	(0.017)
Low Income	Consumption at	0.471	0.579	0.621	0.469	0.340	0.472
	$t-1$ (in logs)	(0.031)	(0.035)	(0.038)	(0.034)	(0.027)	(0.009)
	Treatment	0.112	0.083	0.052	0.039	0.026	0.065
	(2pm - 7pm)	(0.025)	(0.017)	(0.011)	(0.011)	(0.011)	(0.017)
				PCT			
High Income	Consumption at	0.709	0.787	0.810	0.685	0.535	0.676
	$t-1$ (in logs)	(0.033)	(0.027)	(0.024)	(0.028)	(0.026)	(0.008)
	Treatment	-0.103	-0.064	-0.036	-0.044	-0.044	-0.085
	(2pm - 7pm)	(0.033)	(0.025)	(0.016)	(0.018)	(0.023)	(0.014)
Low Income	Consumption at	0.727	0.784	0.802	0.701	0.584	0.687
	$t-1$ (in logs)	(0.036)	(0.030)	(0.029)	(0.033)	(0.032)	(0.007)
	Treatment	-0.082	-0.045	-0.021	-0.016	-0.014	-0.061
	(2pm - 7pm)	(0.031)	(0.022)	(0.013)	(0.012)	(0.017)	(0.015)

TABLE 4.3. *Quantile Mean Group estimator results by Income Levels. MG denotes the standard Mean Group estimator. IHD denotes in-home display and PCT is programmable communicating thermostats. Standard errors are in parentheses.*

		Quantiles					MG
		0.10	0.25	0.50	0.75	0.90	
				Control Group			
Family years	Consumption at	0.493	0.618	0.666	0.515	0.369	0.507
	$t-1$ (in logs)	(0.027)	(0.029)	(0.029)	(0.028)	(0.021)	(0.009)
	Treatment	0.161	0.119	0.062	0.050	0.041	0.099
	(2pm - 7pm)	(0.015)	(0.012)	(0.008)	(0.008)	(0.008)	(0.017)
Young years	Consumption at	0.439	0.535	0.575	0.444	0.337	0.447
	$t-1$ (in logs)	(0.029)	(0.029)	(0.030)	(0.027)	(0.020)	(0.009)
	Treatment	0.135	0.106	0.063	0.047	0.053	0.090
	(2pm - 7pm)	(0.016)	(0.013)	(0.010)	(0.011)	(0.009)	(0.019)
				Portal			
Family years	Consumption at	0.471	0.602	0.654	0.505	0.364	0.502
	$t-1$ (in logs)	(0.027)	(0.030)	(0.032)	(0.029)	(0.020)	(0.009)
	Treatment	0.109	0.086	0.053	0.043	0.032	0.067
	(2pm - 7pm)	(0.017)	(0.017)	(0.015)	(0.016)	(0.021)	(0.017)
Young years	Consumption at	0.467	0.578	0.615	0.471	0.357	0.479
	$t-1$ (in logs)	(0.017)	(0.014)	(0.013)	(0.014)	(0.013)	(0.009)
	Treatment	0.075	0.051	0.026	0.009	-0.011	0.035
	(2pm - 7pm)	(0.028)	(0.019)	(0.014)	(0.015)	(0.020)	(0.018)
				IHD			
Family years	Consumption at	0.508	0.647	0.696	0.540	0.384	0.535
	$t-1$ (in logs)	(0.029)	(0.030)	(0.031)	(0.032)	(0.025)	(0.009)
	Treatment	0.084	0.056	0.028	0.014	0.001	0.033
	(2pm - 7pm)	(0.025)	(0.020)	(0.013)	(0.011)	(0.017)	(0.016)
Young years	Consumption at	0.440	0.527	0.550	0.422	0.327	0.435
	$t-1$ (in logs)	(0.032)	(0.037)	(0.040)	(0.037)	(0.026)	(0.009)
	Treatment	0.117	0.086	0.047	0.034	0.009	0.058
	(2pm - 7pm)	(0.024)	(0.018)	(0.015)	(0.013)	(0.015)	(0.018)
				PCT			
Family years	Consumption at	0.719	0.780	0.801	0.689	0.556	0.679
	$t-1$ (in logs)	(0.037)	(0.030)	(0.026)	(0.032)	(0.032)	(0.007)
	Treatment	-0.078	-0.042	-0.020	-0.018	-0.003	-0.055
	(2pm - 7pm)	(0.036)	(0.025)	(0.015)	(0.016)	(0.022)	(0.014)
Young years	Consumption at	0.718	0.789	0.809	0.696	0.562	0.683
	$t-1$ (in logs)	(0.032)	(0.029)	(0.027)	(0.030)	(0.027)	(0.007)
	Treatment	-0.101	-0.062	-0.034	-0.037	-0.044	-0.083
	(2pm - 7pm)	(0.028)	(0.021)	(0.013)	(0.013)	(0.018)	(0.015)

TABLE 4.4. *Quantile Mean Group estimator results by Family Stages. MG denotes the standard Mean Group estimator. IHD denotes in-home display and PCT is programmable communicating thermostats. Standard errors are in parentheses.*

savings among families with no children who were provided a PCT, and the gains ranges from 3.3% at the 0.5 quantile to 9.6% at the 0.10 quantile. However, PCT does not seem to be an effective technology for middle aged families at the upper quantiles of the conditional distribution of electricity consumption.

5. Conclusions

In this paper, we extend the Common Correlated Effects (CCE) approach to the estimation and inference of dynamic panel quantile regression models with interactive effects. We propose a new quantile estimator and show that it is consistent and asymptotically normal under conditions that are standard in the literature. An important condition is that the individual models need to be augmented by a sufficiently large number of lagged dependent variables. We also show that the approach offers the best finite sample performance in the class of dynamic quantile regression estimators, as long as the time series dimension of the panel is large. Lastly, we demonstrate how the approach can be used in practice by documenting how the use of different technologies that allow consumers to be informed about electricity prices and consumption are associated with energy savings. Using data from a large scale randomized experiment that contains more than 6 million observations, we semiparametrically estimate a dynamic equation for electricity consumption with slope heterogeneity and cross-sectional dependence. The results offer several new insights useful for policy, while illustrating that the average effect does not summarize the distributional effect of the technologies.

Several directions remain to be investigated. Inference procedures are proposed but they require a detailed investigation in the case of long run effects. Moreover, although T is relatively large in our empirical application, offering an estimation approach that helps to reduce potential biases in short T applications seems of fundamental importance. A bias-corrected mean quantile group estimator is being investigated for the case of heterogeneous quantile coefficients following closely the approach developed in Chudik and Pesaran (2015).

References

Allcott, H. (2011): "Rethinking Real-Time Electricity Pricing," *Resource and Energy Economics*, 33, 820–842.

Arellano, M., and S. Bonhomme (2016): "Nonlinear panel data estimation via quantile regressions," *The Econometrics Journal*.

Bai, J. (2009): "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77(4), 1229–1279.

Chernozhukov, V., I. Fernández-Val, J. Hahn, and W. Newey (2013): "Average and Quantile Effects in Nonseparable Panel Models," *Econometrica*, 81, 535–580.

Chernozhukov, V., I. Fernández-Val, S. Hoderlein, H. Holzmann, and W. Newey (2015): "Nonparametric identification in panels using quantiles," *Journal of Econometrics*, 188(2), 378 – 392.

Chudik, A., K. Mohaddes, M. Pesaran, and M. Raissi (2015): "A dynamic panel data approach to the analysis of the effects of government debt and inflation on output growth: Is there a debt threshold effect?," mimeo, USC.

Chudik, A., and M. H. Pesaran (2015): "Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors," *Journal of Econometrics*, 188(2), 393 – 420.

Galvao, A. F. (2011): "Quantile regression for dynamic panel data with fixed effects," *Journal of Econometrics*, 164(1), 142 – 157.

Galvao, A. F., and L. Wang (2015): "Efficient minimum distance estimator for quantile regression fixed effects panel data," *Journal of Multivariate Analysis*, 133(C), 1–26.

Harding, H., and C. Lamarche (2016a): "Penalized Quantile Regression with Semiparametric Correlated Effects: An Application with Heterogeneous Preferences," *Journal of Applied Econometrics, forthcoming.*

Harding, M., and C. Lamarche (2014): "Estimating and testing a quantile regression model with interactive effects," *Journal of Econometrics*, pp. 101 – 113.

Harding, M., and C. Lamarche (2016b): "Empowering Consumers through Smart Technology: Experimental Evidence on the Consequences of Time-of-Use Electricity Pricing," Forthcoming, Journal of Policy Analysis and Management.

Ito, K. (2014): "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing," *American Economic Review*, 104(2), 537063.

Jessoe, K., and D. Rapson (2014): "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, 4(104), 1417–1438.

Kato, K., A. F. Galvao, and G. Montes-Rojas (2012): "Asymptotics for Panel Quantile Regression Models with Individual Effects," *Journal of Econometrics*, 170, 76–91.

Knight, K. (1998): "Limiting Distributions for L_1 Regression Estimators Under General Conditions," *Annals of Statistics*, 26, 755–770.

Koenker, R. (2004): "Quantile Regression for Longitudinal Data," *Journal of Multivariate Analysis*, 91, 74–89.

——— (2005): *Quantile Regression*. Cambridge University Press.

Lamarche, C. (2010): "Robust Penalized Quantile Regression Estimation for Panel Data," *Journal of Econometrics*, 157, 396–408.

Moon, H. R., and M. Weidner (2015a): "Dynamic Linear Panel Regression Models with Interactive Effects," *Econometric Theory*, pp. 1–38.

——— (2015b): "Linear Regression for Panel With Unknown Number of Factors as Interactive Fixed Effects," *Econometrica*, 83(4), 1543–1579.

Pesaran, M. H. (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

Pesaran, M. H., and A. Chudik (2014): "Aggregation in large dynamic panels," *Journal of Econometrics*, 178, Part 2(0), 273 – 285.
Rosen, A. M. (2012): "Set identification via quantile restrictions in short panels," *Journal of Econometrics*, 166(1), 127 – 137.

Appendix A. Mathematical Proofs

Throughout this appendix, we omit, at times, τ in $\pi_i(\tau)$ for notational simplicity. We define $\Delta_i(\pi_i) = \mathbb{M}_i(\pi_i) - \mathbb{M}_i(\pi_{i0})$ and

$$\begin{aligned} \mathbb{M}_i(\pi_i) &:= \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_{it} - \mathbf{X}_{it}'\pi_i\right), \quad \mathbb{M}_i(\pi_{i0}) := \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_{it} - \mathbf{X}_{it}'\pi_{i0}\right), \\ \mathbb{H}_i(\pi_i) &:= \frac{1}{T}\sum_{t=1}^{T}\psi_\tau\left(Y_{it} - \mathbf{X}_{it}'\pi_i\right)\mathbf{X}_{it}, \quad H_i(\pi_i) := E(\mathbb{H}_i(\pi_i)) = E(\tau - F_i(\cdot)|\mathbf{X}_{it}), \\ \mathbf{J}_i &:= \frac{\partial H_i(\pi_i)}{\partial \pi_i} = E(f_i(0|\mathbf{X}_{it})\mathbf{X}_{it}\mathbf{X}_{it}'). \end{aligned}$$

where $\rho_\tau = u(\tau - I(u \leq 0))$ is the quantile regression check function and $\psi_\tau(u) = \tau - I(u \leq 0)$ is the quantile influence function. The proofs refer to Knight's (1998) identity: $\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau + \int_0^v (I(v \leq s) - I(v \leq 0))ds$.

A.1. Proofs of the Main Results

Proof of Theorem 1. For each $\eta > 0$, define the ball $\mathcal{B}_i(\eta) := \{\pi_i : \|\pi_i - \pi_{i0}\|_1 \leq \eta\}$ and the boundary $\partial\mathcal{B}_i(\eta) := \{\pi_i : \|\pi_i - \pi_{i0}\|_1 = \eta\}$. For each $\pi_i \notin \mathcal{B}_i(\eta)$, define $\bar{\pi}_i = r_i\pi_i + (1 - r_i)\pi_{i0}$, where $r_i = \eta/\|\pi_i - \pi_{i0}\|$. Note that $\bar{\pi}_i$ is in the boundary $\partial\mathcal{B}_i(\eta)$. Because the objective function is convex,

$$r_i\left(\mathbb{M}_i(\pi_i) - \mathbb{M}_i(\pi_{i0})\right) \geq \mathbb{M}_i(\bar{\pi}_i) - \mathbb{M}_i(\pi_{i0}) = \mathbb{M}_i(\bar{\pi}_i) = E(\Delta_i(\bar{\pi}_i)) + \left(\mathbb{M}_i(\bar{\pi}_i) - E(\Delta_i(\bar{\pi}_i))\right), \quad \text{(A.1)}$$

Note that $\mathbb{M}_i(\pi_{i0})$ is naturally equal to zero by definition and that $E(\Delta_i(\bar{\pi}_i)) \geq \epsilon_\eta$ for all $1 \leq i \leq N$.

Consider now $\|\hat{\pi}_i - \pi_{i0}\|_1 > \eta$ which implies that $\hat{\pi}_i \notin \mathcal{B}_i(\eta)$ for all $1 \leq i \leq N$. It follows that $\mathbb{M}_i(\hat{\pi}_i) \leq \mathbb{M}_i(\pi_{i0})$ for some $1 \leq i \leq N$ by definition of $\hat{\pi}_i = \arg\min\{\mathbb{M}_i(\pi_i)\}$.

Note that $\hat{\pi}_i \notin \mathcal{B}_i(\eta)$ implies $\hat{\pi}_i \leq 0$ by definition. Thus, by equation (A.1), the following inclusion relationshiop is true:

$$\{\|\hat{\pi}_i - \pi_{i0}\|_1 > \eta\} \subset \left\{\max_{1\leq i\leq N}\sup_{\pi_i\in\mathcal{B}_i(\eta)}|\Delta_i(\pi_i) - E(\Delta_i(\pi_i))| \geq \epsilon_\eta\right\}. \quad \text{(A.2)}$$

We therefore need to show that

$$\max_{1\le i\le N} P\left\{\sup_{\pi_i\in\mathcal{B}_i(\eta)} |\Delta_i(\pi_i) - E(\Delta_i(\pi_i))| \ge \epsilon_\eta\right\} = o(N^{-1}), \tag{A.3}$$

which is similar to equation (A.3) in Kato, Galvao and Montes-Rojas (2012).

Without loss of generality, we restrict all the balls $\mathcal{B}_i(\eta)$ to be equal to $\mathcal{B}(\eta)$ by setting $\pi_{i0} = 0$. Thus, $\mathcal{B}_i(\eta) = \mathcal{B}(\eta)$ for all $1 \le i \le N$. Let $g_\pi(u, \mathbf{X}) = \rho_\tau(u - \mathbf{X}'\pi) - \rho_\tau(u)$. We observe that $|g_\pi(u, \mathbf{X}) - g_{\bar\pi}(u, \mathbf{X})| \le C(1 + M)(\|\pi - \bar\pi\|_1)$, for some universal constant C. Since $\mathcal{B}(\eta)$ is a compact subset in $\mathbb{R}^p$ with $p = 1 + p_x + p_y$, $\exists K$ ℓ_1 balls with center $\pi^{(j)}$ and radius $\epsilon/3\kappa$ where $\kappa := C(1 + M)$.

For each $\pi \in \mathcal{B}(\eta)$, there is $j \in \{1, ..., K\}$ such that,

$$|\Delta(\pi) - E(\Delta(\pi))| \le |\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))| + \frac{2\epsilon}{3}. \tag{A.4}$$

The last inequality follows by a property of $g_\pi(u, \mathbf{X})$. Notice that,

$$\begin{aligned} |\Delta(\pi) - E(\Delta(\pi))| - |\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))| &\le |\Delta(\pi) - E(\Delta(\pi)) - \Delta(\pi^{(j)}) + E(\Delta(\pi^{(j)}))| \\ &\le |\Delta(\pi) - E(\Delta(\pi))| + |\Delta(\pi^{(j)}) + E(\Delta(\pi^{(j)}))| \\ &\le C(1+M)\frac{\epsilon}{3\kappa} + C(1+M)\frac{\epsilon}{3\kappa} = \frac{2}{3}\epsilon. \end{aligned}$$

Therefore, following (A.4), we write,

$$\begin{aligned} P\left(\sup_{\pi\in\mathcal{B}(\eta)} |\Delta(\pi) - E(\Delta(\pi))| > \epsilon\right) &\le \sum_{j=1}^{K} P\left(|\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))| + \frac{2}{3}\epsilon > \epsilon\right) \\ &= \sum_{j=1}^{K} P\left(|\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))| > \frac{\epsilon}{3}\right). \end{aligned}$$

For each $j \in \{1, ..., K\}$, we note that $|\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))|$ satisfies the conditions of Lemma 1. Taking $s = 2\log(N)$ and $q = [\sqrt{T}]$ and using the fact that $\log(N)/\sqrt{T} \to 0$,

$$P\left(|\Delta(\pi^{(j)}) - E(\Delta(\pi^{(j)}))|\right) \le const \times \left(\exp\{-2\log(N)\} + \sqrt{T}Ba^{[\sqrt{T}]}\right),$$

and therefore,

$$P\left\{\sup_{\pi_i\in\mathcal{B}_i(\eta)} |\Delta_i(\pi_i) - E(\Delta_i(\pi_i))| \ge \epsilon_\eta\right\} \le K \times const \times \left(\exp\{-2\log(N)\} + \sqrt{T}Ba^{[\sqrt{T}]}\right). \tag{A.5}$$

Because $(\log(N))^2/T \to 0$, the right hand side of equation (A.5) is $o(N^{-1})$ which leads to the weak consistency result. □

Proof of Theorem 2. Under the stated assumptions, the results follows directly from Theorem 1. By definition, $\hat{\vartheta}(\tau) = N^{-1}\sum_{i=1}^{N}\hat{\vartheta}_i(\tau)$. Thus,

$$\begin{aligned}\hat{\vartheta}(\tau) - \vartheta(\tau) &= \frac{1}{N}\sum_{i=1}^{N}\hat{\vartheta}_i(\tau) - \vartheta(\tau) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\vartheta}_i(\tau) - \vartheta(\tau)\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ (\hat{\pi}_i(\tau) - \pi_i(\tau)) + \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \bar{v}_i(\tau, \mathbf{X}_i) = o_p(1),\end{aligned}$$

where $\bar{v}_i(\tau, \mathbf{X}_{it}) = \pi_i(\tau) - \pi(\tau)$. The first term converges in probability to zero as established in Theorem 1 and the last equality follows by Assumption 5 and 11.b. $\square$

Proof of Theorem 3. By definition, as in Theorem 2, we have

$$\begin{aligned}\hat{\vartheta}(\tau) - \vartheta(\tau) &= \frac{1}{N}\sum_{i=1}^{N}(\hat{\vartheta}_i(\tau) - \vartheta(\tau)) = \frac{1}{N}\sum_{i=1}^{N}((\hat{\vartheta}_i(\tau) - \vartheta_i(\tau)) + (\vartheta_i(\tau) - \vartheta(\tau))) \\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ ((\hat{\pi}_i(\tau) - \pi_i(\tau)) + (\pi_i(\tau) - \pi(\tau)))\end{aligned}$$

It follows that,

$$\sqrt{NT}\left(\hat{\vartheta}(\tau) - \vartheta(\tau)\right) = \frac{\sqrt{N}}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \sqrt{T}\,(\hat{\pi}_i(\tau) - \pi_i(\tau)) + \frac{\sqrt{N}}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \sqrt{T}(\pi_i(\tau) - \pi(\tau)). \quad \text{(A.6)}$$

We now obtain the asymptotic representation of $\hat{\pi}_i(\tau) - \pi_i(\tau)$ following closely Galvao and Wang (2015). We use an expansion of $H_i(\hat{\pi}_i) = E(\mathbb{H}_i(\hat{\pi}_i))$ around π_{i0} to obtain,

$$H_i(\hat{\pi}_i) = H_i(\pi_{i0}) + \mathbf{J}_i(\hat{\pi}_i(\tau) - \pi_{i0}(\tau)) + O_p\left((\hat{\pi}_i(\tau) - \pi_{i0}(\tau))^2\right),$$

where $\mathbf{J}_i := \partial H_i(\pi_i)/\partial \pi_{i0} = E(f_i(0|\mathbf{X}_{it})\mathbf{X}_{it}\mathbf{X}_{it}'))$. Basic manipulations lead to:

$$\begin{aligned}\hat{\pi}_i(\tau) - \pi_{i0}(\tau) &= \mathbf{J}_i^{-1}\left(H_i(\hat{\pi}_i) - H_i(\pi_{i0}) + O_p\left((\hat{\pi}_i(\tau) - \pi_i(\tau))^2\right)\right) \\ &= -\mathbf{J}_i^{-1}\mathbb{H}_i(\pi_{i0}) - \mathbf{J}_i^{-1}\left(\mathbb{H}_i(\hat{\pi}_i) - \mathbb{H}_i(\pi_{i0})\right) - \mathbf{J}_i^{-1}\left(H_i(\hat{\pi}_i) - H_i(\pi_{i0})\right) \\ &\quad + \mathbf{J}_i^{-1}\left(\mathbb{H}_i(\hat{\pi}_i)\right) + \mathbf{J}_i^{-1}O_p\left((\hat{\pi}_i(\tau) - \pi_i(\tau))^2\right) \\ &= -\mathbf{J}_i^{-1}\mathbb{H}_i(\pi_{i0}) - \mathbf{J}_i^{-1}\left[(\mathbb{H}_i(\hat{\pi}_i) - \mathbb{H}_i(\pi_{i0})) - (H_i(\hat{\pi}_i) - H_i(\pi_{i0}))\right] \\ &\quad + \mathbf{J}_i^{-1}\left(\mathbb{H}_i(\hat{\pi}_i)\right) + \mathbf{J}_i^{-1}O_p\left((\hat{\pi}_i(\tau) - \pi_i(\tau))^2\right).\end{aligned}$$

For fixed N, the second term in the last expression is $o_p(1)$. In the case of panel data with individual parameters, we need to find the order of $\max_{1\leq i\leq N}\left[(\mathbb{H}_i(\hat{\pi}_i) - \mathbb{H}_i(\pi_{i0})) - (H_i(\hat{\pi}_i) - H_i(\pi_{i0}))\right]$.

Lemma 2 establishes that order. Therefore for each i, we have that,

$$\hat{\pi}_i(\tau) - \pi_i(\tau) = -\mathbf{J}_i^{-1}\mathbb{H}_i(\pi_{i0}) + O(d_N) + O(T^{-1}) + \mathbf{J}_i^{-1}O_p\left((\hat{\pi}_i(\tau) - \pi_i(\tau))^2\right), \tag{A.7}$$

where $d_N := T^{-(1-c)}\log(N) \vee T^{-1/2}\delta_N^{1/2}(\log(N))^{1/2}$ and $\delta_N = \sqrt{\log(N)/T}$ when $|\log(\delta_N)| \asymp \log(N)$. Substituting equation (A.7) in equation (A.6), after basic simplifications, we obtain

$$\begin{aligned}\sqrt{NT}\left(\hat{\vartheta}(\tau) - \vartheta(\tau)\right) &= \frac{\sqrt{N}}{TN}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \sqrt{T}\mathbf{J}_i^{-1}\sum_{t=1}^{T}\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi)\mathbf{X}_{it} + \sqrt{TN}O(d_N) \\ &\quad + \frac{\sqrt{N}}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \sqrt{T}(\pi_i(\tau) - \pi(\tau)). \end{aligned} \tag{A.8}$$

The second term is equation (A.8) $O_p((NT)^{1/2}d_N)$. Using the d_N rate implied by Lemma 2 for a sufficiently small c, we have that $(NT)^{1/2}d_N = N^{1/2}\log(N)^{1/4}T^{-1/4}\log(N)^{1/2}$. Therefore, if $N^2(\log(N))^3/T \to 0$, $\|\hat{\vartheta}(\tau) - \vartheta(\tau)\| = O_p((NT)^{-1/2})$ if the last term is of the same other than the first term in the expression.

We apply similar arguments to obtain an asymptotic representation of $\pi_i(\tau) - \pi(\tau)$. Note that if $\hat{\pi}_i(\tau) = \pi_i(\tau)$, we still have that $\|\hat{\vartheta}(\tau) - \vartheta(\tau)\| = O_p((NT)^{-1/2})$ although the first term in equation (A.8) is zero. If $\hat{\pi}_i(\tau) \neq \pi_i(\tau)$, we require some efforts. Similarly as for $\hat{\pi}_i(\tau) - \pi_{i0}(\tau)$, we write $\pi_i(\tau) - \pi(\tau) = \mathbf{J}_i^{-1}(\mathbb{H}_i(\pi_i) - \mathbb{H}_i(\pi)) + \mathbf{J}_i^{-1}O_p\left((\pi_i(\tau) - \pi(\tau))^2\right)$. Note that the second term is bounded by Assumption 5 and Assumption 11. Therefore, the last term in equation (A.8) can be written as,

$$\frac{\sqrt{N}}{N}\sum_{i=1}^{N}\mathbf{\Xi}_i \circ \frac{\sqrt{T}}{T}\mathbf{J}_i^{-1}\sum_{t=1}^{T}\mathbf{X}_{it}[\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i) - \psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi)],$$

or alternatively as,

$$\frac{1}{\sqrt{TN}}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{\Xi}_i \circ \mathbf{J}_i^{-1}\mathbf{X}_{it}[I(Y_{it} \leq \mathbf{X}_{it}'\pi) - I(Y_{it} \leq \mathbf{X}_{it}'\pi_i)],$$

which has mean zero and bounded variance by Assumptions 5 and 11.b.

By standard arguments, as N and T tends to infinity under the conditions on p_T and the relative rate of N and T, $\sqrt{NT}\left(\hat{\vartheta}(\tau) - \vartheta(\tau)\right) \rightsquigarrow \mathcal{N}(\mathbf{0}, \mathbf{V}_\psi + \mathbf{V}_v)$. $\square$

A.2. **Lemmas**

In recent years, there has been considerable progress on establishing the rate of the remainder terms of the Bahadur representation of the quantile regression estimator. The next two lemmas are used in the proofs of Theorems 1, 2, and 3.

Lemma 1 (Corollary C.1, Kato et al. (2012))**.** *Let f be a function on $\mathcal{S}$ a Polish space and let $\{\xi_t : t \geq 1\}$ be a stationary process taking values in a measurable space $(\mathcal{S}, \mathcal{K})$. Assume that $\mathcal{K}$ is a Borel σ-field. Let $sup_{x\in\mathcal{X}}|f(\xi_t)| \leq U$ for some constant U and $E(f(\xi_t)) = 0$. Take $q \in [1, T/2]$. Then,*

$$P\left(|\sum_{t=1}^{T} f(\xi_t)| \geq const \times \left\{\sqrt{(s \vee 1)T}\sigma_q(f) + sqU\right\}\right) \leq 2e^{-s} + 2r\beta(q),$$

where $r := [T/2q]$, $\beta(\cdot)$ denote mixing coefficients, and

$$\sigma_q(f) := Var(f(\xi_t)) + 2\sum_{j=1}^{q-1}(1 - j/q)Cov(f(\xi_1), f(\xi_{1+j})).$$

Lemma 2 (Galvao and Wang (2015))**.** *Under regularity conditions 7-10, for any $c \in (0, 1)$ and δ_N such that $|\log(\delta_N)| \asymp \log(N)$,*

$$\max_{1\leq i\leq N}\{(\mathbb{M}_i(\hat{\pi}_i) - \mathbb{M}_i(\pi_{i0})) - (M_i(\hat{\pi}_i) - M_i(\pi_{i0}))\} = O_p(T^{-(1-c)}\log(N) \vee T^{-1/2}\delta_N^{1/2}(log(N))^{1/2}),$$

and $\max_{1\leq i\leq N}\|\hat{\pi}_i(\tau) - \pi_i(\tau)\| = O_p(\delta_N)$ where $\delta_N = \sqrt{\log(N)/T}$.

Proof. Details of the proof are omitted as the proof is an application of Lemma 1 as discussed in Kato et al. (2012). The results follow immediately by verifying that conditions 7-10 imply conditions A2-A5 and B1-B3 in Lemmas 7 and 8 in Galvao and Wang (2015). □

Supplement: Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Quantile Regression Models

Matthew Harding[‡], Carlos Lamarche[§], and M. Hashem Pesaran[¶]

In this Supplement, we first offer a derivation of the variance of $\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i)$ in the case of dependence. We next extend the simulation results of the paper by offering results on the unfeasible QMG estimator which uses the latent factors, $\mathbf{f}_t$. This extension is presented in the second section of the Supplement. Lastly, we present additional results on the power of the QMG estimator.

S.1. Estimation of Asymptotic Covariance Matrix

Let $\bar{\xi}_i(\tau) := T^{-1/2}\sum_{t=1}^{T}\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i)$. By definition, we have that

$$\mathrm{Var}(\bar{\xi}_i(\tau)) = \frac{1}{T}\sum_{t=1}^{T}\mathrm{Var}(\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i)) + 2\sum_{t\neq t'}^{T}\mathrm{Cov}(\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i), \psi_\tau(Y_{it'} - \mathbf{X}_{it'}'\pi_i)).$$

Note that,

$$\begin{aligned}
\mathrm{Cov}(\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i), \psi_\tau(Y_{it'} - \mathbf{X}_{it'}'\pi_i)) &= \mathrm{Cov}(\tau - I(Y_{it} < \mathbf{X}_{it}'\pi_i), \tau - I(Y_{it'} < \mathbf{X}_{it'}'\pi_i)) \\
&= E((\tau - I(Y_{it} < \mathbf{X}_{it}'\pi_i))(\tau - I(Y_{it'} < \mathbf{X}_{it'}'\pi_i))) \\
&\quad - E(\tau - I(Y_{it} < \mathbf{X}_{it}'\pi_i))E(\tau - I(Y_{it'} < \mathbf{X}_{it'}'\pi_i)) \\
&= E((\tau - I(Y_{it} < \mathbf{X}_{it}'\pi_i))(\tau - I(Y_{it'} < \mathbf{X}_{it'}'\pi_i))),
\end{aligned}$$

because $\tau = F(\mathbf{X}_{it}'\pi_i) = F(\mathbf{X}_{it'}'\pi_i)$ under Assumption 7. It follows that,

$$\begin{aligned}
\mathrm{Cov}(\psi_\tau(Y_{it} - \mathbf{X}_{it}'\pi_i), \psi_\tau(Y_{it'} - \mathbf{X}_{it'}'\pi_i)) &= \tau^2 - \tau E(I(Y_{it} < \mathbf{X}_{it}'\pi_i)) - \tau E(I(Y_{it'} < \mathbf{X}_{it'}'\pi_i)) \\
&\quad + E(I(Y_{it} < \mathbf{X}_{it}'\pi_i)I(Y_{it} < \mathbf{X}_{it}'\pi_i)) \\
&= E(I(Y_{it} < \mathbf{X}_{it}'\pi_i)I(Y_{it} < \mathbf{X}_{it}'\pi_i)) - \tau^2.
\end{aligned}$$

[‡]Department of Economics, University of California Irvine, SSPB 3207, Irvine, CA 9269; Email: harding1@uci.edu

[§]Department of Economics, University of Kentucky, 223G Gatton College of Business and Economics, Lexington, KY 40506-0034; Phone: (859) 257 3371; Email: clamarche@uky.edu

[¶]Department of Economics, University of Southern California, Kaprielian Hall 300, Los Angeles, CA 90089-0253; Email: pesaran@dornsife.usc.edu

Thus, we have that

$$\begin{aligned}\sigma_\psi^2(q) := \mathrm{Var}(\bar{\xi}_i(\tau)) &= \tau(1-\tau) + 2\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)\mathrm{Cov}(\psi_\tau(Y_{i1}-\mathbf{X}_{i1}'\pi_i), \psi_\tau(Y_{i1+j}-\mathbf{X}_{i1+j}'\pi_i)) \\ &= \tau(1-\tau) + 2\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)\left[E(I(Y_{i1}<\mathbf{X}_{i1}'\pi_i)I(Y_{i1+j}<\mathbf{X}_{i1+j}'\pi_i)) - \tau^2\right] \\ &= \tau(1-\tau) + 2\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)\left[E(I(Y_{i1}<\mathbf{X}_{i1}'\pi_i, Y_{i1+j}<\mathbf{X}_{i1+j}'\pi_i)) - \tau^2\right],\end{aligned}$$

as in Section 2.2. It should be noted that $q = 1$ gives $\sigma_\psi^2(1) = \tau(1-\tau)$, the variance of $\psi_\tau = \tau - I(u < 0)$ in the case of independent observations.

S.2. Simulation Evidence

This section offers additional Monte Carlo evidence on the finite sample performance of the proposed estimator for different values of λ. It then compares the performance of the feasible QMG estimator with an unfeasible version of it that uses unknown factors $\mathbf{f}$.

S.2.1. Autoregressive Models

Tables S.1 presents the bias and root mean square error (RMSE) for the slope parameter β in the location shift model with $\lambda \in \{0.25, 0.75\}$, which are different to the value $\lambda = 0.5$ used in the first tables of Section 3. The table presents results for Designs 1-4, showing results for quantile regression estimators at two quantiles, $\tau \in \{0.25, 0.50\}$. The table compares the performance of the quantile regression estimator for a dynamic panel data model (Galvao 2011) denoted by DQR and the quantile mean group (QMG) estimator for a model with interactive effects. The DQR estimator uses y_{it-2} as an instrument for y_{it-1}. The proposed quantile mean group estimator, QMG, is obtained as the cross sectional average of $\hat{\beta}_i(\tau)$ using $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{y}_{t-1}, \bar{x}_t)'$. The sample size is based on $N = 100$ and $T = 200$. Table S.1 shows that he QMG method for β performs extremely well with negligible biases and low RMSE for all values of λ.

Tables S.2 and S.3 report results for λ and θ. While Table S.2 presents simulation results for the case of $\lambda = 0.25$, Table S.3 presents simulation results for the case of $\lambda = 0.75$. In terms of relative performance performance between DQR and QMG estimators, the tables do not offer new insights. The evidence continues to suggest that the QMG estimator performs better in small samples than the DQR estimator with considerable gains in terms of bias and MSE. We find however that the absolute bias of the QMG estimator tends to increase as λ increases. The results however are not presented in terms of percentage bias since θ increase as λ increases.

S.2.2. Estimation of Models with Known f

This section compares the results of the estimator QMG defined in Section 2 with the results obtained by employing an unfeasible version of the estimator. The unfeasible estimator is defined as:

$$\tilde{\pi}_i(\tau) = \arg\min_{\pi_i \in \mathbf{\Pi}_i} \sum_{t=1}^{T} \rho_\tau \left(y_{it} - \lambda_i y_{it-1} - \beta_i' \mathbf{x}_{it} - \gamma_i' \mathbf{f}_t \right)$$

Therefore, $\tilde{\pi}_i(\tau)$ is an estimator based on a cross-sectional quantile regression with latent factors, $\mathbf{f}_t$. Moreover, we define $\tilde{\vartheta}(\tau) = \frac{1}{N}\sum_{i=1}^{N} \tilde{\vartheta}_i(\tau) = \frac{1}{N}\sum_{i=1}^{N} (\mathbf{\Xi}_i \circ \tilde{\pi}_i(\tau))$, where $\circ$ denotes Hadamard product and $\mathbf{\Xi}_i = (\iota_i', \mathbf{0}_i')'$ with ι_i denoting a vector of ones. In what follows, we denote this estimator as UNF.

Tables S.4 to Table S.5 present the bias and root mean square error (RMSE) of the QMG and UNF estimators for the slope parameter β in the location shift model with $\lambda = 0.5$. Table S.4 presents results for Designs 1 and 2 and Table S.5 presents results for Designs 3 and 4. The tables show results for quantile regression estimators at two quantiles, $\tau \in \{0.25, 0.50\}$, based on sample sizes of $N \in \{100, 200\}$ and $T \in \{50, 100, 200\}$. The tables show that, as expected, the UNF estimator offers smaller bias and smaller RMSE than its feasible version QMG. We also observe that these differences tend to dissapear as long as both N and T increase.

In the next four tables, Tables S.6 to Table S.9, we present results the bias and root mean square error (RMSE) of the QMG and UNF estimators for the parameters λ and θ. The results continues to indicate that the unfeasible version improves the performance of the feasible version, although again as in the case of the slope parameter β, the finite sample performance of the QMG estimator approximates very closely to the performance of the UNF estimator when $T > 100$.

Naturally, the large-N approximation of the factors would not be needed if the factors $\mathbf{f}_t$ are known. Therefore, the quantiles of the error term (2.12) would be identical of the quantiles of the conditional quantile function (2.4). Given that N can be small in applications, it is natural to investigate using simulations results how well the approximation works for the second and third term in equation (2.12) to be negligible. In other words, the conditional quantiles $Q_{u_{it}+V_{it}}(\tau) \approx Q_{u_{it}}(\tau)$ for large N but they might differ for small N creating biases in the performance of the feasible QMG estimator.

Figure S.1 offers results for the comparison between the distribution of the error term u in equation (2.2) and the distribution of $u+V$ in equation (2.12) when $N \in \{10, 100, 200\}$ is small and $T = 50$. We follow Design 2 where $\mathbf{x}_{it}$ is generated as in equation (3.5) and u distributed as t_4 and χ_1^2 distribution with mean zero. Consistent with the theoretical findings, we find that the differences between distributions dissapear when N increases and for $N > 100$ the differences in the quantiles tends to be negligible.

S.2.3. Power

This section reports additional simulation results for the power of the QMG estimator. In light of the theoretical results, we limit our investigation to document the shape of the power function as the number of time series T units increase. We generate data using Design 1 considering $N = 100$ and $T \in \{100, 200, 400\}$ for the case of Gaussian and χ_3^2 error term. The evidence is presented in Figures S.2 and S.3 at the end of the Supplement.

λ	Design		Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
0.25	1	Bias	0.302	-0.007	0.302	-0.007	0.289	-0.006	0.289	-0.007	0.231	-0.008	0.215	-0.002
0.25	1	RMSE	0.310	0.010	0.309	0.010	0.297	0.008	0.297	0.010	0.240	0.011	0.223	0.005
0.25	2	Bias	0.315	-0.007	0.313	-0.007	0.298	-0.005	0.298	-0.007	0.238	-0.008	0.222	-0.002
0.25	2	RMSE	0.322	0.010	0.321	0.011	0.305	0.008	0.306	0.010	0.247	0.010	0.230	0.005
0.25	3	Bias	0.314	-0.008	0.310	-0.008	0.282	-0.007	0.280	-0.008	0.193	-0.009	0.177	-0.003
0.25	3	RMSE	0.321	0.011	0.318	0.011	0.290	0.010	0.289	0.011	0.200	0.012	0.185	0.005
0.25	4	Bias	0.309	-0.008	0.304	-0.008	0.286	-0.007	0.284	-0.008	0.194	-0.009	0.178	-0.002
0.25	4	RMSE	0.315	0.012	0.310	0.012	0.294	0.009	0.293	0.011	0.201	0.012	0.185	0.005
0.75	1	Bias	0.077	-0.007	0.077	-0.007	0.075	-0.007	0.075	-0.009	0.064	-0.014	0.059	-0.007
0.75	1	RMSE	0.079	0.008	0.079	0.008	0.078	0.008	0.078	0.009	0.068	0.015	0.063	0.007
0.75	2	Bias	0.081	-0.008	0.080	-0.008	0.078	-0.007	0.078	-0.0088	0.066	-0.015	0.062	-0.007
0.75	2	RMSE	0.084	0.009	0.083	0.009	0.080	0.008	0.081	0.010	0.070	0.015	0.065	0.007
0.75	3	Bias	0.081	-0.009	0.080	-0.009	0.075	-0.009	0.075	-0.010	0.052	-0.014	0.048	-0.007
0.75	3	RMSE	0.084	0.010	0.083	0.010	0.078	0.009	0.078	0.011	0.055	0.015	0.051	0.007
0.75	4	Bias	0.076	-0.009	0.076	-0.010	0.079	-0.009	0.078	-0.008	0.052	-0.013	0.047	-0.006
0.75	4	RMSE	0.079	0.009	0.079	0.011	0.081	0.009	0.080	0.009	0.054	0.014	0.050	0.007

TABLE S.1. *Bias and root mean square error (RMSE) of quantile regression estimators for* β *when* $\lambda \in \{0.25, 0.75\}$, $N = 100$ *and* $T = 200$.

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
λ	Design		Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
					Normal Distribution					
0.25	1	Bias	-0.267	0.006	0.313	-0.001	-0.266	0.005	0.313	-0.002
0.25	1	RMSE	0.281	0.015	0.324	0.014	0.281	0.014	0.325	0.015
0.25	2	Bias	-0.292	-0.010	0.300	-0.022	-0.291	-0.010	0.298	-0.022
0.25	2	RMSE	0.305	0.016	0.313	0.026	0.304	0.017	0.313	0.027
0.25	3	Bias	-0.280	0.009	0.326	0.000	-0.264	0.011	0.285	0.004
0.25	3	RMSE	0.295	0.015	0.339	0.014	0.279	0.017	0.298	0.016
0.25	4	Bias	-0.273	0.007	0.321	-0.002	-0.255	0.011	0.280	0.004
0.25	4	RMSE	0.286	0.015	0.332	0.015	0.268	0.018	0.292	0.016
					t_4 distribution					
0.25	1	Bias	-0.247	0.006	0.308	0.000	-0.246	0.008	0.310	0.001
0.25	1	RMSE	0.264	0.013	0.321	0.014	0.263	0.016	0.324	0.017
0.25	2	Bias	-0.247	0.013	0.321	0.010	-0.246	0.015	0.325	0.011
0.25	2	RMSE	0.263	0.017	0.334	0.018	0.262	0.021	0.338	0.020
0.25	3	Bias	-0.239	0.007	0.300	-0.001	-0.225	0.014	0.262	0.007
0.25	3	RMSE	0.257	0.015	0.313	0.016	0.246	0.019	0.274	0.017
0.25	4	Bias	-0.239	0.014	0.314	0.009	-0.225	0.021	0.275	0.017
0.25	4	RMSE	0.258	0.019	0.327	0.018	0.246	0.025	0.287	0.023
					χ^2_3 distribution					
0.25	1	Bias	-0.178	0.011	0.257	0.003	-0.163	0.002	0.235	0.001
0.25	1	RMSE	0.198	0.026	0.271	0.029	0.182	0.015	0.248	0.019
0.25	2	Bias	-0.212	-0.015	0.224	-0.032	-0.196	-0.024	0.204	-0.033
0.25	2	RMSE	0.229	0.028	0.242	0.043	0.213	0.028	0.220	0.038
0.25	3	Bias	-0.194	0.008	0.225	-0.005	-0.159	0.006	0.187	0.003
0.25	3	RMSE	0.214	0.024	0.241	0.029	0.180	0.015	0.202	0.018
0.25	4	Bias	-0.170	0.043	0.276	0.040	-0.131	0.037	0.243	0.046
0.25	4	RMSE	0.193	0.048	0.290	0.048	0.156	0.040	0.256	0.050

TABLE S.2. *Bias and root mean square error (RMSE) of quantile regression estimators when $\lambda = 0.25$, $N = 100$ and $T = 200$.*

λ	Design		$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
						Normal Distribution				
0.75	1	Bias	-0.107	0.012	1.180	-0.029	-0.106	0.012	1.184	-0.019
0.75	1	RMSE	0.130	0.018	1.223	0.057	0.129	0.017	1.229	0.055
0.75	2	Bias	-0.130	-0.003	1.170	-0.085	-0.129	-0.003	1.150	-0.079
0.75	2	RMSE	0.150	0.012	1.221	0.101	0.148	0.013	1.200	0.097
0.75	3	Bias	-0.118	0.016	1.249	-0.029	-0.100	0.019	1.205	-0.005
0.75	3	RMSE	0.141	0.020	1.300	0.061	0.125	0.023	1.260	0.057
0.75	4	Bias	-0.111	0.014	1.225	-0.030	-0.091	0.018	1.177	0.000
0.75	4	RMSE	0.132	0.018	1.267	0.062	0.116	0.022	1.217	0.056
						t_4 distribution				
0.75	1	Bias	-0.099	0.014	1.177	-0.024	-0.098	0.016	1.192	-0.020
0.75	1	RMSE	0.124	0.018	1.224	0.058	0.124	0.022	1.244	0.067
0.75	2	Bias	-0.096	0.021	1.229	0.008	-0.096	0.023	1.245	0.010
0.75	2	RMSE	0.123	0.024	1.274	0.054	0.123	0.028	1.294	0.064
0.75	3	Bias	-0.100	0.016	1.175	-0.032	-0.083	0.024	1.138	-0.003
0.75	3	RMSE	0.126	0.021	1.224	0.061	0.113	0.028	1.184	0.056
0.75	4	Bias	-0.097	0.023	1.229	-0.004	-0.079	0.031	1.193	0.026
0.75	4	RMSE	0.125	0.027	1.276	0.053	0.112	0.035	1.236	0.060
						χ^2_3 distribution				
0.75	1	Bias	-0.071	0.029	1.028	-0.046	-0.064	0.013	0.935	-0.020
0.75	1	RMSE	0.106	0.037	1.081	0.105	0.096	0.020	0.983	0.063
0.75	2	Bias	-0.102	0.002	0.941	-0.149	-0.095	-0.014	0.853	-0.123
0.75	2	RMSE	0.128	0.023	1.003	0.177	0.119	0.021	0.908	0.136
0.75	3	Bias	-0.079	0.028	0.930	-0.081	-0.051	0.016	0.832	-0.020
0.75	3	RMSE	0.109	0.036	0.985	0.129	0.084	0.022	0.879	0.068
0.75	4	Bias	-0.045	0.062	1.085	0.059	-0.014	0.048	0.995	0.115
0.75	4	RMSE	0.088	0.067	1.134	0.105	0.070	0.050	1.038	0.130

TABLE S.3. *Bias and root mean square error (RMSE) of quantile regression estimators when $\lambda = 0.75$, $N = 100$ and $T = 200$.*

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
Design 1: Location shift with homogeneous slopes														
100	50	Bias	-0.031	-0.052	-0.031	-0.053	-0.031	-0.050	-0.035	-0.056	-0.049	-0.066	-0.025	-0.035
100	50	RMSE	0.032	0.054	0.033	0.055	0.032	0.051	0.037	0.057	0.050	0.068	0.026	0.037
100	100	Bias	-0.015	-0.024	-0.015	-0.024	-0.015	-0.022	-0.017	-0.025	-0.023	-0.029	-0.011	-0.014
100	100	RMSE	0.017	0.025	0.016	0.025	0.017	0.024	0.019	0.027	0.025	0.030	0.013	0.015
100	200	Bias	-0.008	-0.008	-0.007	-0.008	-0.007	-0.007	-0.009	-0.009	-0.012	-0.013	-0.005	-0.004
100	200	RMSE	0.009	0.010	0.008	0.010	0.008	0.009	0.010	0.011	0.013	0.014	0.007	0.006
200	50	Bias	-0.032	-0.057	-0.031	-0.056	-0.032	-0.053	-0.037	-0.060	-0.048	-0.068	-0.025	-0.037
200	50	RMSE	0.032	0.058	0.032	0.057	0.033	0.054	0.037	0.061	0.049	0.069	0.026	0.037
200	100	Bias	-0.015	-0.026	-0.015	-0.025	-0.015	-0.023	-0.017	-0.026	-0.024	-0.032	-0.012	-0.015
200	100	RMSE	0.016	0.026	0.016	0.026	0.016	0.024	0.018	0.027	0.024	0.032	0.012	0.016
200	200	Bias	-0.008	-0.011	-0.007	-0.011	-0.007	-0.010	-0.008	-0.011	-0.012	-0.015	-0.006	-0.006
200	200	RMSE	0.008	0.012	0.008	0.011	0.008	0.011	0.009	0.012	0.012	0.015	0.006	0.007
Design 2: Location shift with heterogeneous slopes														
100	50	Bias	-0.031	-0.053	-0.031	-0.053	-0.031	-0.049	-0.035	-0.055	-0.049	-0.066	-0.025	-0.035
100	50	RMSE	0.033	0.054	0.033	0.055	0.032	0.051	0.037	0.057	0.050	0.068	0.026	0.036
100	100	Bias	-0.015	-0.023	-0.014	-0.023	-0.015	-0.022	-0.017	-0.025	-0.023	-0.029	-0.011	-0.014
100	100	RMSE	0.017	0.025	0.016	0.024	0.017	0.023	0.019	0.026	0.025	0.030	0.013	0.015
100	200	Bias	-0.008	-0.009	-0.008	-0.009	-0.007	-0.007	-0.009	-0.009	-0.012	-0.012	-0.005	-0.004
100	200	RMSE	0.009	0.010	0.009	0.011	0.008	0.009	0.010	0.011	0.013	0.014	0.007	0.006
200	50	Bias	-0.032	-0.057	-0.032	-0.056	-0.033	-0.054	-0.037	-0.061	-0.048	-0.068	-0.024	-0.036
200	50	RMSE	0.033	0.058	0.032	0.057	0.033	0.054	0.038	0.061	0.049	0.069	0.025	0.037
200	100	Bias	-0.015	-0.025	-0.015	-0.025	-0.016	-0.024	-0.017	-0.027	-0.023	-0.032	-0.011	-0.015
200	100	RMSE	0.016	0.026	0.015	0.026	0.016	0.024	0.018	0.027	0.024	0.032	0.012	0.016
200	200	Bias	-0.007	-0.011	-0.007	-0.010	-0.008	-0.010	-0.009	-0.011	-0.012	-0.015	-0.006	-0.007
200	200	RMSE	0.008	0.012	0.008	0.011	0.008	0.011	0.009	0.012	0.012	0.015	0.006	0.007

TABLE S.4. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for β in Designs 1 and 2. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
Design 3: Location-scale shift with homogeneous slopes														
100	50	Bias	-0.036	-0.059	-0.035	-0.059	-0.036	-0.056	-0.040	-0.063	-0.045	-0.062	-0.022	-0.033
100	50	RMSE	0.038	0.061	0.037	0.061	0.037	0.058	0.042	0.065	0.046	0.064	0.024	0.035
100	100	Bias	-0.017	-0.026	-0.017	-0.026	-0.017	-0.024	-0.019	-0.027	-0.023	-0.030	-0.011	-0.013
100	100	RMSE	0.019	0.027	0.018	0.027	0.018	0.025	0.021	0.029	0.024	0.031	0.012	0.015
100	200	Bias	-0.008	-0.010	-0.008	-0.010	-0.008	-0.009	-0.009	-0.011	-0.011	-0.013	-0.005	-0.005
100	200	RMSE	0.010	0.012	0.010	0.012	0.009	0.010	0.010	0.013	0.012	0.014	0.006	0.007
200	50	Bias	-0.037	-0.063	-0.036	-0.063	-0.036	-0.058	-0.041	-0.065	-0.045	-0.064	-0.022	-0.033
200	50	RMSE	0.037	0.064	0.037	0.064	0.037	0.059	0.042	0.066	0.046	0.065	0.023	0.034
200	100	Bias	-0.018	-0.029	-0.017	-0.028	-0.017	-0.026	-0.019	-0.029	-0.022	-0.029	-0.010	-0.014
200	100	RMSE	0.018	0.030	0.018	0.029	0.018	0.027	0.020	0.030	0.022	0.030	0.011	0.014
200	200	Bias	-0.008	-0.012	-0.008	-0.012	-0.008	-0.011	-0.009	-0.013	-0.011	-0.014	-0.005	-0.006
200	200	RMSE	0.009	0.013	0.009	0.013	0.009	0.012	0.010	0.013	0.012	0.014	0.006	0.007
Design 4: Location-scale shift with heterogeneous slopes														
100	50	Bias	-0.037	-0.062	-0.036	-0.060	-0.035	-0.057	-0.040	-0.064	-0.047	-0.065	-0.023	-0.034
100	50	RMSE	0.038	0.063	0.038	0.062	0.037	0.059	0.043	0.066	0.048	0.067	0.024	0.036
100	100	Bias	-0.017	-0.026	-0.017	-0.026	-0.017	-0.023	-0.020	-0.028	-0.022	-0.029	-0.011	-0.013
100	100	RMSE	0.019	0.027	0.019	0.028	0.018	0.025	0.021	0.029	0.024	0.030	0.012	0.015
100	200	Bias	-0.008	-0.010	-0.008	-0.010	-0.008	-0.009	-0.009	-0.011	-0.010	-0.012	-0.005	-0.004
100	200	RMSE	0.009	0.012	0.010	0.012	0.009	0.010	0.010	0.012	0.012	0.014	0.006	0.006
200	50	Bias	-0.035	-0.062	-0.036	-0.063	-0.037	-0.059	-0.041	-0.066	-0.045	-0.065	-0.022	-0.034
200	50	RMSE	0.036	0.063	0.037	0.064	0.038	0.060	0.042	0.067	0.046	0.065	0.023	0.035
200	100	Bias	-0.018	-0.029	-0.017	-0.029	-0.017	-0.026	-0.019	-0.029	-0.022	-0.030	-0.010	-0.014
200	100	RMSE	0.018	0.029	0.018	0.029	0.018	0.026	0.019	0.029	0.022	0.030	0.011	0.015
200	200	Bias	-0.008	-0.012	-0.008	-0.013	-0.008	-0.011	-0.009	-0.013	-0.011	-0.014	-0.005	-0.006
200	200	RMSE	0.009	0.013	0.009	0.013	0.009	0.012	0.010	0.014	0.012	0.015	0.006	0.007

TABLE S.5. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for β in Designs 3 and 4. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
			Normal Distribution							
100	50	Bias	0.029	0.051	-0.028	-0.043	0.032	0.053	-0.017	-0.028
100	50	RMSE	0.040	0.058	0.059	0.069	0.044	0.062	0.057	0.068
100	100	Bias	0.019	0.029	-0.011	-0.017	0.017	0.028	-0.009	-0.013
100	100	RMSE	0.024	0.033	0.031	0.036	0.024	0.034	0.034	0.036
100	200	Bias	0.009	0.009	-0.007	-0.005	0.008	0.009	-0.007	-0.005
100	200	RMSE	0.015	0.017	0.022	0.022	0.014	0.016	0.023	0.024
200	50	Bias	0.034	0.058	-0.021	-0.044	0.035	0.058	-0.010	-0.030
200	50	RMSE	0.038	0.062	0.043	0.060	0.040	0.061	0.040	0.053
200	100	Bias	0.018	0.030	-0.012	-0.021	0.017	0.030	-0.008	-0.014
200	100	RMSE	0.021	0.032	0.023	0.029	0.022	0.033	0.024	0.028
200	200	Bias	0.010	0.014	-0.004	-0.008	0.009	0.013	-0.003	-0.005
200	200	RMSE	0.013	0.016	0.016	0.017	0.013	0.016	0.016	0.018
			t_4 distribution							
100	50	Bias	0.032	0.049	-0.025	-0.042	0.037	0.055	-0.017	-0.031
100	50	RMSE	0.044	0.058	0.061	0.075	0.051	0.067	0.063	0.073
100	100	Bias	0.019	0.027	-0.013	-0.017	0.019	0.029	-0.015	-0.018
100	100	RMSE	0.025	0.032	0.033	0.039	0.028	0.036	0.039	0.042
100	200	Bias	0.010	0.010	-0.004	-0.003	0.012	0.012	-0.005	-0.002
100	200	RMSE	0.015	0.015	0.020	0.023	0.018	0.019	0.025	0.027
200	50	Bias	0.035	0.054	-0.025	-0.046	0.039	0.061	-0.021	-0.036
200	50	RMSE	0.040	0.058	0.046	0.064	0.046	0.065	0.051	0.082
200	100	Bias	0.018	0.028	-0.014	-0.020	0.019	0.029	-0.013	-0.021
200	100	RMSE	0.021	0.031	0.027	0.032	0.024	0.033	0.031	0.037
200	200	Bias	0.009	0.012	-0.007	-0.008	0.011	0.014	-0.005	-0.007
200	200	RMSE	0.012	0.015	0.018	0.019	0.015	0.017	0.020	0.022
			χ^2_3 distribution							
100	50	Bias	0.055	0.075	-0.023	-0.035	0.029	0.041	-0.011	-0.015
100	50	RMSE	0.076	0.092	0.102	0.111	0.047	0.058	0.074	0.084
100	100	Bias	0.031	0.037	-0.009	-0.015	0.018	0.017	0.001	-0.007
100	100	RMSE	0.045	0.049	0.061	0.059	0.028	0.027	0.039	0.040
100	200	Bias	0.017	0.018	-0.003	-0.003	0.009	0.006	-0.001	0.000
100	200	RMSE	0.030	0.030	0.046	0.044	0.017	0.016	0.029	0.027
200	50	Bias	0.051	0.072	-0.034	-0.050	0.028	0.039	-0.015	-0.026
200	50	RMSE	0.063	0.082	0.077	0.088	0.038	0.049	0.053	0.061
200	100	Bias	0.029	0.038	-0.015	-0.022	0.015	0.019	-0.006	-0.008
200	100	RMSE	0.038	0.044	0.046	0.049	0.022	0.026	0.031	0.034
200	200	Bias	0.017	0.020	-0.005	-0.007	0.009	0.009	-0.001	-0.002
200	200	RMSE	0.023	0.026	0.029	0.030	0.014	0.014	0.020	0.021

TABLE S.6. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for λ and θ in Design 1. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
						Normal Distribution				
100	50	Bias	0.041	0.063	-0.003	-0.018	0.040	0.061	0.003	-0.007
100	50	RMSE	0.049	0.069	0.048	0.058	0.049	0.069	0.055	0.066
100	100	Bias	0.026	0.034	0.005	-0.001	0.024	0.033	0.008	0.004
100	100	RMSE	0.031	0.039	0.031	0.033	0.030	0.039	0.032	0.036
100	200	Bias	-0.008	-0.007	-0.041	-0.039	-0.008	-0.007	-0.040	-0.038
100	200	RMSE	0.014	0.015	0.046	0.045	0.014	0.015	0.046	0.044
200	50	Bias	0.040	0.063	-0.009	-0.034	0.040	0.063	-0.001	-0.021
200	50	RMSE	0.044	0.066	0.039	0.053	0.044	0.066	0.037	0.048
200	100	Bias	0.026	0.038	0.006	-0.003	0.025	0.037	0.009	0.002
200	100	RMSE	0.028	0.039	0.021	0.023	0.028	0.039	0.026	0.026
200	200	Bias	0.008	0.012	-0.007	-0.011	0.007	0.010	-0.007	-0.009
200	200	RMSE	0.011	0.014	0.016	0.019	0.011	0.014	0.017	0.019
						t_4 distribution				
100	50	Bias	0.045	0.060	-0.001	-0.019	0.049	0.066	0.007	-0.008
100	50	RMSE	0.053	0.068	0.054	0.065	0.060	0.076	0.061	0.067
100	100	Bias	0.032	0.039	0.014	0.011	0.032	0.042	0.013	0.009
100	100	RMSE	0.036	0.043	0.033	0.036	0.038	0.047	0.038	0.039
100	200	Bias	-0.006	-0.006	-0.035	-0.033	-0.004	-0.003	-0.035	-0.032
100	200	RMSE	0.012	0.013	0.040	0.040	0.014	0.015	0.044	0.042
200	50	Bias	0.029	0.047	-0.037	-0.058	0.033	0.054	-0.034	-0.070
200	50	RMSE	0.035	0.052	0.053	0.072	0.041	0.059	0.058	0.306
200	100	Bias	0.016	0.025	-0.018	-0.023	0.017	0.027	-0.017	-0.025
200	100	RMSE	0.020	0.029	0.029	0.034	0.023	0.031	0.033	0.039
200	200	Bias	-0.014	-0.011	-0.052	-0.053	-0.012	-0.009	-0.051	-0.053
200	200	RMSE	0.016	0.014	0.055	0.056	0.016	0.014	0.054	0.056
						χ^2_3 distribution				
100	50	Bias	0.067	0.087	-0.003	-0.009	0.041	0.053	0.013	0.008
100	50	RMSE	0.084	0.102	0.101	0.123	0.055	0.066	0.073	0.082
100	100	Bias	0.028	0.034	-0.015	-0.020	0.015	0.015	-0.005	-0.012
100	100	RMSE	0.043	0.046	0.062	0.061	0.027	0.026	0.040	0.042
100	200	Bias	0.012	0.013	-0.012	-0.012	0.003	0.001	-0.011	-0.010
100	200	RMSE	0.027	0.027	0.048	0.046	0.016	0.015	0.031	0.029
200	50	Bias	0.059	0.076	-0.020	-0.036	0.035	0.045	0.000	-0.012
200	50	RMSE	0.068	0.084	0.070	0.109	0.042	0.053	0.046	0.055
200	100	Bias	0.031	0.041	-0.009	-0.018	0.017	0.021	0.000	-0.003
200	100	RMSE	0.039	0.047	0.047	0.048	0.022	0.026	0.027	0.028
200	200	Bias	0.034	0.039	0.032	0.030	0.027	0.028	0.036	0.035
200	200	RMSE	0.037	0.042	0.043	0.043	0.029	0.030	0.041	0.040

TABLE S.7. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for λ and θ in Design 2. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
			Normal Distribution							
100	50	Bias	0.039	0.058	-0.023	-0.048	0.042	0.066	-0.004	-0.011
100	50	RMSE	0.048	0.066	0.060	0.076	0.051	0.073	0.062	0.067
100	100	Bias	0.023	0.032	-0.007	-0.015	0.024	0.035	0.001	0.000
100	100	RMSE	0.028	0.037	0.032	0.036	0.030	0.041	0.034	0.039
100	200	Bias	0.011	0.013	-0.004	-0.005	0.011	0.016	-0.002	0.005
100	200	RMSE	0.017	0.018	0.022	0.022	0.018	0.021	0.023	0.024
200	50	Bias	0.039	0.065	-0.026	-0.050	0.044	0.074	-0.003	-0.012
200	50	RMSE	0.044	0.068	0.044	0.063	0.049	0.077	0.040	0.047
200	100	Bias	0.021	0.033	-0.013	-0.024	0.023	0.038	-0.003	-0.006
200	100	RMSE	0.024	0.036	0.027	0.038	0.027	0.040	0.026	0.027
200	200	Bias	0.010	0.014	-0.007	-0.010	0.011	0.017	-0.001	0.000
200	200	RMSE	0.013	0.017	0.017	0.019	0.014	0.020	0.016	0.016
			t_4 distribution							
100	50	Bias	0.037	0.054	-0.032	-0.056	0.044	0.069	-0.012	-0.005
100	50	RMSE	0.049	0.063	0.067	0.083	0.057	0.079	0.069	0.223
100	100	Bias	0.022	0.029	-0.010	-0.017	0.028	0.040	-0.001	0.002
100	100	RMSE	0.028	0.035	0.036	0.042	0.034	0.046	0.035	0.041
100	200	Bias	0.012	0.012	-0.004	-0.005	0.015	0.019	0.002	0.008
100	200	RMSE	0.017	0.018	0.023	0.024	0.019	0.024	0.022	0.025
200	50	Bias	0.036	0.058	-0.033	-0.051	0.048	0.074	-0.010	-0.014
200	50	RMSE	0.041	0.063	0.051	0.068	0.054	0.079	0.049	0.080
200	100	Bias	0.019	0.030	-0.017	-0.025	0.025	0.039	-0.006	-0.008
200	100	RMSE	0.023	0.033	0.028	0.035	0.029	0.042	0.028	0.029
200	200	Bias	0.010	0.014	-0.007	-0.009	0.014	0.019	0.000	0.000
200	200	RMSE	0.013	0.016	0.018	0.020	0.017	0.022	0.018	0.020
			χ^2_3 distribution							
100	50	Bias	0.060	0.082	-0.036	-0.055	0.041	0.060	0.010	0.008
100	50	RMSE	0.082	0.100	0.116	0.126	0.055	0.072	0.076	0.080
100	100	Bias	0.037	0.044	-0.016	-0.026	0.018	0.025	-0.002	0.003
100	100	RMSE	0.049	0.055	0.067	0.069	0.030	0.036	0.045	0.050
100	200	Bias	0.016	0.017	-0.010	-0.015	0.009	0.010	-0.002	0.002
100	200	RMSE	0.028	0.028	0.044	0.046	0.017	0.018	0.028	0.028
200	50	Bias	0.059	0.082	-0.041	-0.063	0.038	0.057	0.002	0.001
200	50	RMSE	0.068	0.089	0.076	0.090	0.046	0.064	0.047	0.052
200	100	Bias	0.031	0.041	-0.022	-0.032	0.017	0.024	-0.002	-0.002
200	100	RMSE	0.038	0.047	0.048	0.054	0.023	0.029	0.031	0.033
200	200	Bias	0.016	0.020	-0.011	-0.015	0.010	0.012	0.000	0.002
200	200	RMSE	0.024	0.026	0.034	0.035	0.015	0.017	0.021	0.022

TABLE S.8. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for λ and θ in Design 3. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ		Parameter: λ		Parameter: θ	
			UNF	QMG	UNF	QMG	UNF	QMG	UNF	QMG
						Normal Distribution				
100	50	Bias	0.045	0.067	-0.012	-0.040	0.049	0.076	0.009	0.001
100	50	RMSE	0.052	0.071	0.052	0.065	0.056	0.081	0.053	0.061
100	100	Bias	0.016	0.026	-0.020	-0.023	0.021	0.033	-0.006	-0.005
100	100	RMSE	0.023	0.032	0.036	0.042	0.027	0.039	0.036	0.036
100	200	Bias	0.007	0.010	-0.009	-0.008	0.010	0.014	-0.004	0.003
100	200	RMSE	0.013	0.016	0.022	0.024	0.016	0.020	0.023	0.024
200	50	Bias	0.043	0.068	-0.013	-0.038	0.049	0.077	0.007	-0.005
200	50	RMSE	0.047	0.071	0.038	0.057	0.052	0.080	0.036	0.044
200	100	Bias	0.013	0.025	-0.027	-0.037	0.015	0.029	-0.017	-0.021
200	100	RMSE	0.017	0.028	0.035	0.043	0.020	0.032	0.030	0.033
200	200	Bias	0.009	0.013	-0.009	-0.012	0.011	0.017	-0.003	-0.001
200	200	RMSE	0.012	0.016	0.018	0.020	0.014	0.020	0.016	0.016
						t_4 distribution				
100	50	Bias	0.048	0.066	-0.009	-0.033	0.061	0.087	0.016	0.007
100	50	RMSE	0.055	0.074	0.055	0.071	0.069	0.094	0.070	0.076
100	100	Bias	0.014	0.022	-0.023	-0.028	0.020	0.033	-0.015	-0.011
100	100	RMSE	0.023	0.029	0.041	0.045	0.028	0.040	0.038	0.041
100	200	Bias	0.019	0.019	0.012	0.011	0.022	0.027	0.017	0.024
100	200	RMSE	0.022	0.023	0.025	0.026	0.026	0.030	0.028	0.033
200	50	Bias	0.030	0.050	-0.046	-0.067	0.043	0.070	-0.017	-0.030
200	50	RMSE	0.036	0.055	0.060	0.079	0.048	0.073	0.050	0.057
200	100	Bias	0.023	0.034	-0.007	-0.014	0.027	0.040	0.002	0.000
200	100	RMSE	0.027	0.037	0.025	0.030	0.032	0.044	0.029	0.031
200	200	Bias	0.011	0.015	-0.005	-0.007	0.014	0.020	-0.001	0.001
200	200	RMSE	0.014	0.018	0.016	0.019	0.017	0.023	0.018	0.018
						χ^2_3 distribution				
100	50	Bias	0.073	0.100	-0.023	-0.041	0.050	0.072	0.022	0.026
100	50	RMSE	0.089	0.114	0.098	0.104	0.062	0.083	0.071	0.077
100	100	Bias	0.074	0.082	0.060	0.051	0.058	0.064	0.076	0.080
100	100	RMSE	0.083	0.089	0.089	0.081	0.063	0.069	0.088	0.092
100	200	Bias	0.058	0.059	0.072	0.068	0.048	0.050	0.079	0.084
100	200	RMSE	0.062	0.063	0.084	0.080	0.051	0.053	0.084	0.089
200	50	Bias	0.066	0.091	-0.024	-0.049	0.045	0.063	0.015	0.012
200	50	RMSE	0.077	0.099	0.085	0.089	0.053	0.070	0.055	0.060
200	100	Bias	0.038	0.049	-0.008	-0.018	0.024	0.031	0.012	0.011
200	100	RMSE	0.043	0.054	0.042	0.046	0.029	0.035	0.033	0.033
200	200	Bias	0.025	0.029	0.007	0.002	0.018	0.021	0.017	0.019
200	200	RMSE	0.029	0.033	0.028	0.028	0.021	0.023	0.026	0.027

TABLE S.9. *Bias and root mean square error (RMSE) of the unfeasible QMG estimator and the feasible QMG estimator for λ and θ in Design 4. In all the variations of the model, $\lambda = 0.5$. UNF denotes the unfeasible version of the estimator.*

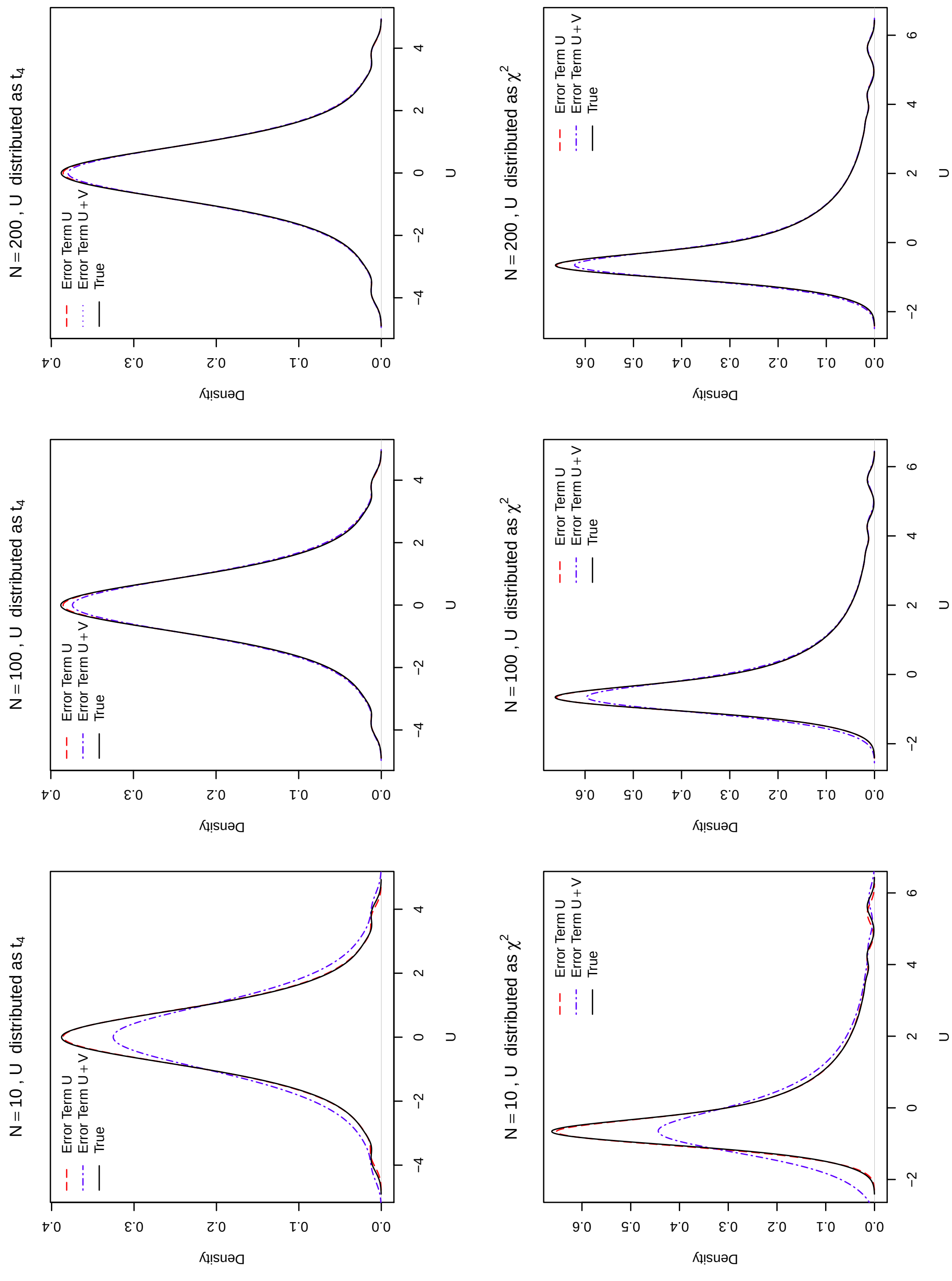

Figure S.1. *Monte Carlo results for the comparison between the distribution of U and the distribution of $U{+}V$ when N is small.*

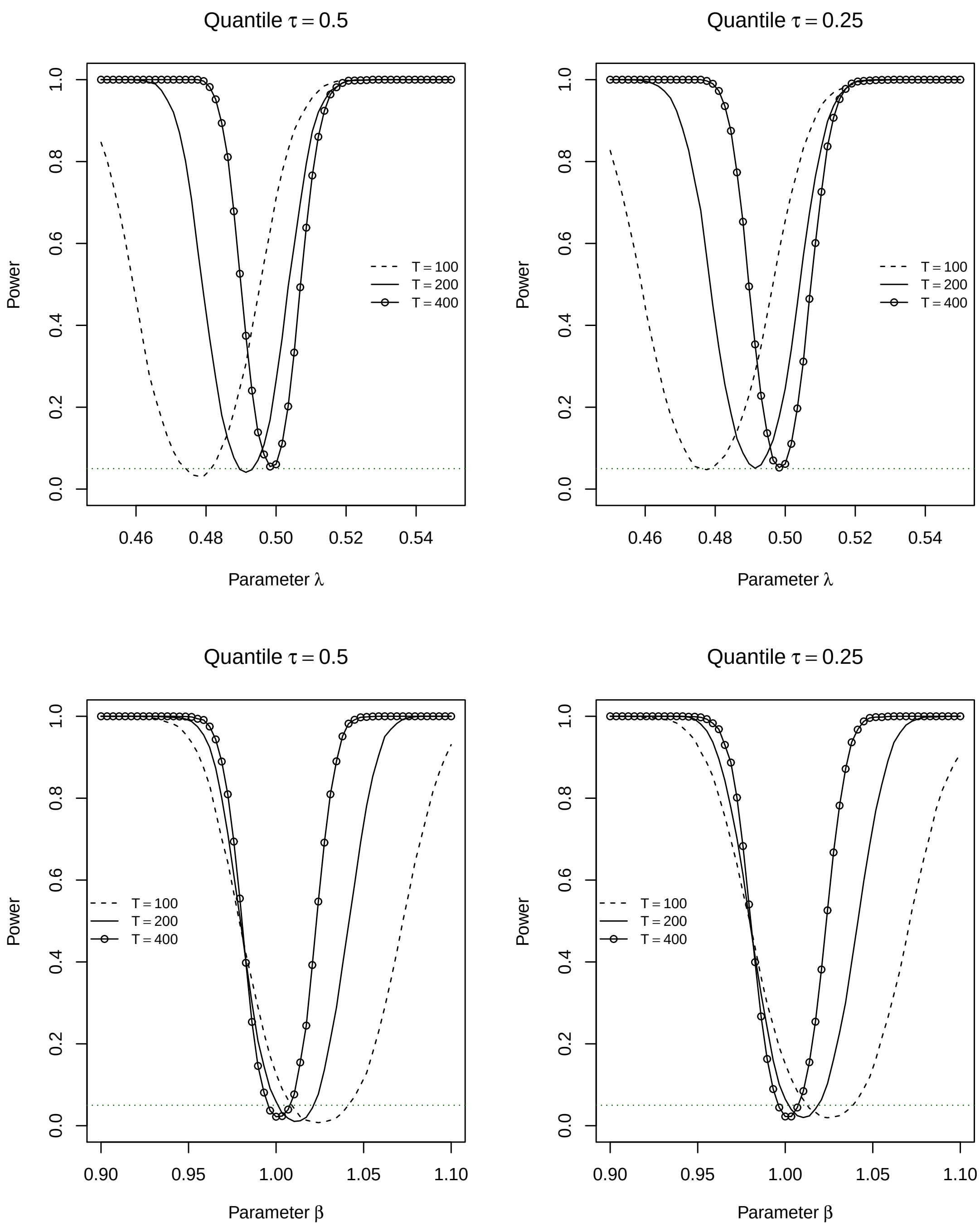

FIGURE S.2. *Power of the QMG estimator against different alternatives. The figures shows results when* $u \sim \mathcal{N}(0,1)$.

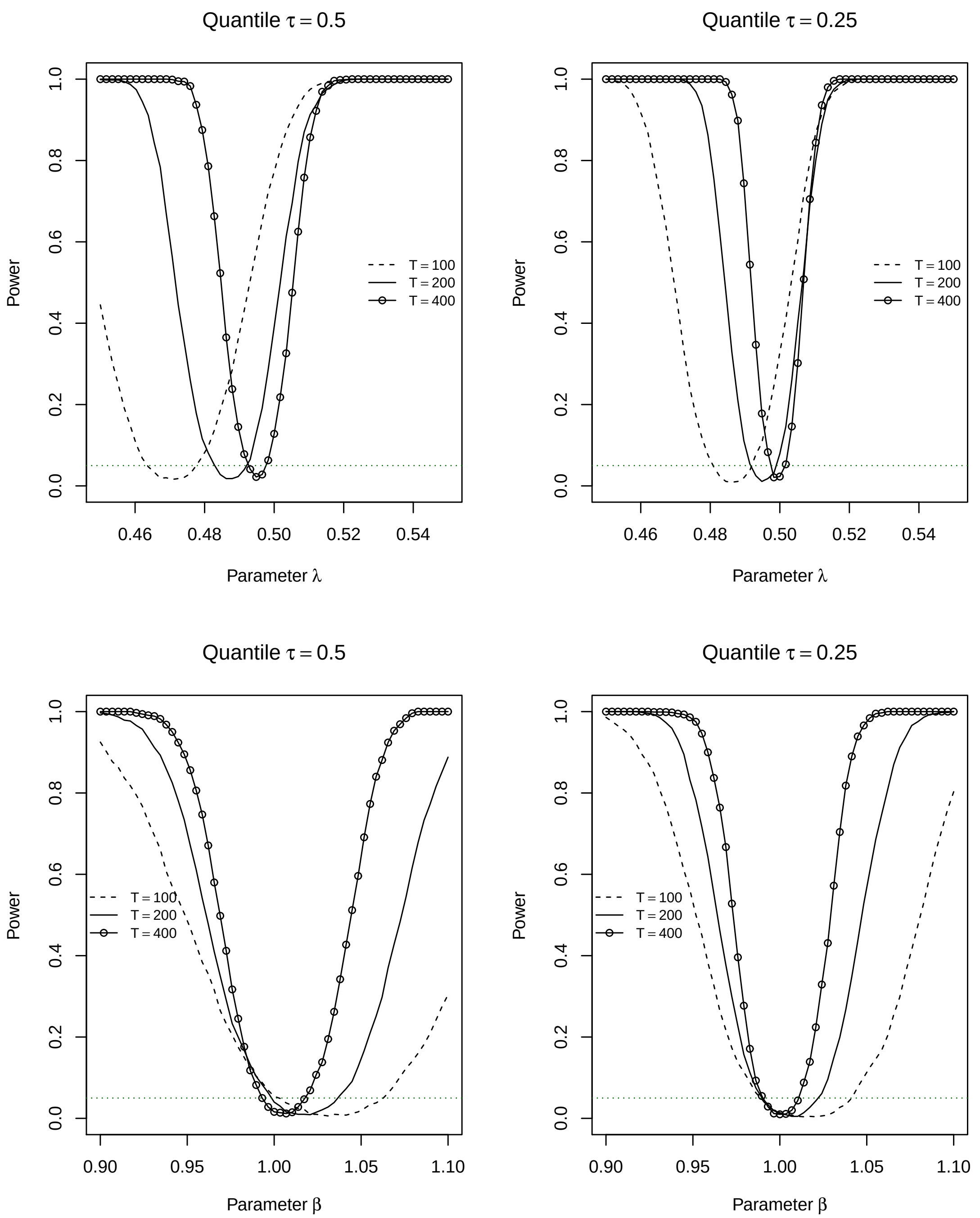

FIGURE S.3. *Power of the QMG estimator against different alternatives. The figures shows results when* $u \sim \chi^2_3$.