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

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

December 17, 2019

In this Appendix, we present mathematical derivations and supporting theoretical results (Section S.1), and additional simulation evidence (Section S.2).

S.1. Mathematical derivations

S.1.1. Definitions

The proofs make use of Knight's (1998) identity: $\rho_\tau(u-v)-\rho_\tau(u) = -v\psi_\tau + \int_0^v (I(v \leq s) - I(v \leq 0))ds$, 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.

Throughout this appendix, we omit, at times, τ in $\boldsymbol{\pi}_i(\tau)$ for notational simplicity. Recall that $\boldsymbol{\pi}_i := (\lambda_i, \boldsymbol{\beta}_i', \alpha_i, \boldsymbol{\delta}_i')'$ where $\boldsymbol{\delta}_i := (\boldsymbol{\delta}_{i1}', \boldsymbol{\delta}_{i2}', \ldots, \boldsymbol{\delta}_{ip_T}')'$ and $\boldsymbol{\theta}_i := (\lambda_i, \boldsymbol{\beta}_i', \alpha_i, \boldsymbol{\gamma}_i')'$. Note that $\boldsymbol{\pi}_i$ is a $(2+p_x)+(1+p_T)(1+p_x)$ dimensional vector and $\boldsymbol{\theta}_i$ is a $2+p_x+r$ dimensional vector. Also, recall that $\mathbf{X}_{it} = (y_{it-1}, \mathbf{x}_{it}', 1, \bar{\mathbf{z}}_t', \bar{\mathbf{z}}_{t-1}', ..., \bar{\mathbf{z}}_{t-p_T}')'$, $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{\mathbf{x}}_t')'$, $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, $\bar{\mathbf{x}}_t = N^{-1}\sum_{i=1}^N \mathbf{x}_{it}$, and $\mathbf{W}_{it} = (y_{it-1}, \mathbf{x}_{it}', 1, \mathbf{f}_t')'$.

Consider the following models:

$$y_{it} = \mathbf{W}_{it}'\boldsymbol{\theta}_i + \xi_{it} \qquad \text{(S.1.1)}$$

$$y_{it} = \mathbf{X}_{it}'\boldsymbol{\pi}_i + h_{it,N} + \xi_{it}, \qquad \text{(S.1.2)}$$

*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, and Trinity College, Cambridge; Kaprielian Hall 300, Los Angeles, CA 90089-0253; Email: pesaran@usc.edu

where

$$h_{it,N} = \sum_{l=p_T+1}^{\infty} \bar{\mathbf{z}}'_{t-l}\boldsymbol{\delta}_{il} + O_p(N^{-1/2}). \tag{S.1.3}$$

S.1.2. **Technical Conditions:**

We consider the following regularity conditions:

Assumption S.1. *There exist a series of constants independent of i and τ such that $\sup_{i,\tau} \|\boldsymbol{\gamma}_i(\tau)\| < K_{\boldsymbol{\gamma}}$, $\sup_i \|y_{i0}\| < K_y$, and additionally a constant M exists such that $\sup_{i,t} \|\boldsymbol{W}_{it}\| \leq M$ (a.s.).*

Assumption S.2. *Let G_i be the conditional distribution of $u_{it}(\tau)$ as in equation (2.3) and assume that the conditional densities g_i are continuous, uniformly bounded away from 0 and ∞, with continuous derivatives everywhere. Moreover, for each $\eta > 0$,*

$$\epsilon_\eta := \inf_i \inf_{\|\boldsymbol{\theta}\|_1=\eta} E\left[\int_0^{\boldsymbol{W}'_{i1}\boldsymbol{\theta}} \left(G_i(s|\boldsymbol{W}_{i1}) - \tau\right) ds\right].$$

Assumption S.3. *Let $\mathbf{J}_i := E[g_i(0|\mathbf{W}_{it})\mathbf{W}_{it}\mathbf{W}'_{it}]$, $\mathbf{D}_i := T^{-1/2}\sum_{t=1}^{T}\psi_\tau(y_{it} - \mathbf{W}'_{it}\boldsymbol{\theta}_0)\mathbf{W}_{it}$, $\dot{\boldsymbol{\Xi}}_i := \boldsymbol{\Xi}_i\boldsymbol{\Xi}'_i$, and $\mathbf{V}_i := Var(\mathbf{D}_i)$. The following conditions hold:*

(a) *Let $\mathbf{J}_N = N^{-1}\sum_{i=1}^{N}\dot{\boldsymbol{\Xi}}_i \circ \mathbf{J}_i$ and $\mathbf{V}_N = N^{-1}\sum_{i=1}^{N}\dot{\boldsymbol{\Xi}}_i \circ \mathbf{V}_i$. The limit $\mathbf{J} := \lim_{N\to\infty}\mathbf{J}_N$, $\mathbf{V} := \lim_{N\to\infty}\mathbf{V}_N$, and $\mathbf{V}_\psi := \lim_{N\to\infty}\mathbf{J}_N^{-1}\mathbf{V}_N\mathbf{J}_N^{-1}$ exist and are non-singular.*

(b) *Let $\mathbf{V}_v := Var(\boldsymbol{\vartheta}_i(\tau))$. The limiting form of $\mathbf{V}_v$ exist and is non-singular.*

Similar conditions are used in the literature (see, e.g., Kato, Galvao and Montes-Rojas (2012), and Galvao, Lamarche and Lima (2013)). Assumption S.1 is needed for the consistency of the estimator and for obtaining a well-defined limiting distribution. It requires that the regressors are strictly bounded, with the implication that the support of the error distributions is bounded and all coefficients, including the factor loadings, are bounded too. Assumption S.2 is an identification condition and is similar to Assumptions (A3), (A1) and (B1) in Kato, Galvao and Montes-Rojas (2012). Assumption S.3 has two parts which correspond to the case of heterogeneous and homogeneous coefficients. The first part is standard in the panel quantile literature for models with homogeneous coefficients 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 S.3.b allows for slope heterogeneity while guaranteeing that the covariance matrix of the QMG estimator is well defined.

S.1.3. **Proofs**

This section provides a proof using a number of high level assumptions. Further work is required to develop a rigorous asymptotic theory when a growing number of variables is used to approximate latent factors in a dynamic quantile regression model. See Remarks S.1 and S.2 below.

Proof of Theorem 1. The proof is divided in two parts. First, we show uniform consistency of the proposed estimator by demonstrating that the feasible and infeasible optimization problems are equivalent as N, T and $p_T \to \infty$. The second part of the proof establishes consistency of $\hat{\boldsymbol{\theta}}_i(\tau)$. Under the conditions of Proposition S.1, the consistency of $\hat{\boldsymbol{\theta}}_i(\tau)$ implies the consistency of the first $(1+p_x)$ elements of the reduced form coefficients, $\hat{\boldsymbol{\pi}}_i(\tau)$.

[*Part 1: Asymptotic equivalence of objective functions*] For each i, define

$$
\begin{aligned}
Q_{i,\infty}(\tau,\boldsymbol{\theta}_i) &= E\left[\rho_\tau(y_{it}-\mathbf{W}_{it}'\boldsymbol{\theta}_i)\right] \\
Q_{i,T}(\tau,\boldsymbol{\theta}_i) &= \frac{1}{T}\sum_{t=1}^{T}\rho_\tau(y_{it}-\mathbf{W}_{it}'\boldsymbol{\theta}_i) \\
\hat{Q}_{i,T,N}(\tau,\boldsymbol{\pi}_i) &= \frac{1}{T}\sum_{t=1}^{T}\rho_\tau(y_{it}-\mathbf{X}_{it}'\boldsymbol{\pi}_i).
\end{aligned}
$$

We establish the required result in two steps. First, we prove that $Q_{i,T}(\tau,\boldsymbol{\theta}_i)$ converges uniformly to $Q_{i,\infty}(\tau,\boldsymbol{\theta}_i)$ in $\boldsymbol{\theta}_i$ and τ. Second, we show that the difference between the feasible optimization problem that uses $\hat{Q}_{i,T,N}(\tau,\boldsymbol{\pi}_i)$ and the infeasible $Q_{i,T}(\tau,\boldsymbol{\theta}_i)$ converges to zero as N, T and $p_T \to \infty$. It follows then that $(\hat{\lambda}_i, \hat{\boldsymbol{\beta}}_i')$ as the solution of $\min \hat{Q}_{i,T,N}(\tau,\boldsymbol{\pi}_i)$ converges to the solution of $\min_{\boldsymbol{\theta}_i} Q_{i,\infty}(\tau,\boldsymbol{\theta}_i)$.

The first step is to show that,

$$
\sup_{i,\tau,\boldsymbol{\theta}} |Q_{i,T}(\tau,\boldsymbol{\theta}_i) - Q_{i,\infty}(\tau,\boldsymbol{\theta}_i)| = o_p(1). \tag{S.1.4}
$$

Note that $\boldsymbol{\theta} \mapsto \rho_\tau(y-\mathbf{W}'\boldsymbol{\theta})$ is continuous for y and $\mathbf{W}$. Moreover, the dominating function corresponding to the quantile regression check function exists under Assumptions 5, S.1, and S.2 and it is equal to $\rho_\tau(y-\mathbf{W}'\boldsymbol{\theta}) \leq K(|\alpha| + |\lambda||y_{-1}| + \|\mathbf{x}\|_1\|\boldsymbol{\beta}\|_1 + \|\boldsymbol{\gamma}\|_1\|\mathbf{f}\|_1)$ for some constant $K > 0$. Then, we can conclude that (S.1.4) holds by Lemma 2.4 of Newey and McFadden (1994).

The second part of the proof uses a version of Knight's (1998) identity: $|\rho_\tau(u-v) - \rho_\tau(u)| \leq 3|v|$. We begin by noticing that by equations (S.1.1) and (S.1.2), we can write

$$\begin{aligned}\sup_i \left|\hat{Q}_{i,T,N}(\tau,\cdot) - Q_{i,T}(\tau,\cdot)\right| &= \sup_i \left|\frac{1}{T}\sum_{t=1}^{T} \rho_\tau(\xi_{it} + h_{it,N}) - \rho_\tau(\xi_{it})\right| \leq K\frac{1}{T}\sum_{t=1}^{T}\sup_i |h_{it,N}| \\ &\leq K\frac{1}{T}\sum_{t=1}^{T}\sup_i \left|\sum_{l=p_T+1}^{\infty} \bar{\mathbf{z}}'_{t-l}\boldsymbol{\delta}_{il}\right| + O_p\left(\frac{1}{\sqrt{N}}\right).\end{aligned}$$

The last inequality is obtained by using the definition of $h_{it,N}$ in equation (S.1.3). Under Assumption 7, $\sup_i \|\boldsymbol{\delta}_{il}\| < K\rho^l$ for all i and l, then the first term can be bounded by

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\sup_i \left|\sum_{l=p_T+1}^{\infty} \bar{\mathbf{z}}'_{t-l}\boldsymbol{\delta}_{il}\right| &\leq K\rho^{p_T+1}\sum_{j=0}^{\infty}\rho^j\left(\frac{1}{T}\sum_{t=1}^{T}\|\bar{\mathbf{z}}_{t-j-p_T-1}\|\right) \\ &\leq \left(\frac{K\rho^{p_T+1}}{1-\rho}\right)\sup_j\left(\frac{1}{T}\sum_{t=1}^{T}\|\bar{\mathbf{z}}_{t-j-p_T-1}\|\right).\end{aligned}$$

Under Assumptions 4, 5 and 7, and by the conditions in Proposition S.1, it follows that

$$\sup_i \left|\hat{Q}_{i,T,N}(\tau,\cdot) - Q_{i,T}(\tau,\cdot)\right| \leq \left(\frac{K\rho^{p_T+1}}{1-\rho}\right)\sup_j\left(\frac{1}{T}\sum_{t=1}^{T}\|\bar{\mathbf{z}}_{t-j-p_T-1}\|\right) + O_p\left(\frac{1}{\sqrt{N}}\right),$$

which tends to zero as N, T, and $p_T \to \infty$.

Remark S.1. The rate $p_T^3/T \to \varkappa$, $0 < \varkappa < \infty$ in Proposition S.1 below guarantees that the approach developed in Chudik and Pesaran (2015) is consistent. Using cross-sectional averages and their p_T lagged values requires to balance two properties: (1) when p_T is large, we can approximate $\mathbf{f}_t$ with $\bar{\mathbf{z}}_t$ and its lagged values, and (2) the rate at which p_T raises with T is sufficiently restrictive to ensure that individual estimates of $\hat{\boldsymbol{\pi}}_i(\tau)$ are consistent. The implication is that the number of regressors are not too many relative to T.

Remark S.2. The issue of the rate at which p_T raises with T is similar to the result established in He and Shao (2000)'s Corollary 2.1, where they establish consistency and asymptotic normality for convex loss function with finitely many jump discontinuities, but they do not allow for the panel aspect of our problem and the actual sample size is T and not $T - p_T$ as in our dynamic case. (See also their Example 2 on the spatial median where

$p_x^2/T \to 0$ is needed). The rates differ from the one needed in least squares regressions where the objective function is differentiable everywhere, and it is required that p^3/T tends to a bounded constant as in Chudik and Pesaran (2015).

The second part of our argument is that since the two objective functions are asymptotically equivalent, we work directly with the infeasible estimator by considering $\mathbf{W}_{it}$ that replaces $\mathbf{X}_{it}$. The development of the proof below follows closely Kato, Galvao and Montes-Rojas (2012) and Galvao and Wang (2015).

[*Part 2: Consistency of quantile coefficients*] For each $\eta > 0$, define the ball $\mathcal{B}_i(\eta) := \{\boldsymbol{\theta}_i : \|\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i0}\|_1 \leq \eta\}$ and the boundary $\partial\mathcal{B}_i(\eta) := \{\boldsymbol{\theta}_i : \|\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i0}\|_1 = \eta\}$. For each $\boldsymbol{\theta}_i \notin \mathcal{B}_i(\eta)$, define $\bar{\boldsymbol{\theta}}_i = r_i\boldsymbol{\theta}_i + (1-r_i)\boldsymbol{\theta}_{i0}$, where $r_i = \eta/\|\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i0}\|$. Note that $\bar{\boldsymbol{\theta}}_i$ is in the boundary $\partial\mathcal{B}_i(\eta)$. Because the objective function is convex,

$$\begin{aligned} r_i\left(Q_{i,T}(\boldsymbol{\theta}_i) - Q_{i,T}(\boldsymbol{\theta}_{i0})\right) &\geq Q_{i,T}(\bar{\boldsymbol{\theta}}_i) - Q_{i,T}(\boldsymbol{\theta}_{i0}) = Q_{i,T}(\bar{\boldsymbol{\theta}}_i) \\ &= E(\Delta_i(\bar{\boldsymbol{\theta}}_i)) + \left(Q_{i,T}(\bar{\boldsymbol{\theta}}_i) - E(\Delta_i(\bar{\boldsymbol{\theta}}_i))\right), \end{aligned} \tag{S.1.5}$$

where $\Delta_i(\bar{\boldsymbol{\theta}}_i) = Q_{i,T}(\bar{\boldsymbol{\theta}}_i) - Q_{i,T}(\boldsymbol{\theta}_{i0})$ and that $E(\Delta_i(\bar{\boldsymbol{\theta}}_i)) \geq \epsilon_\eta$ for all $1 \leq i \leq N$.

Consider now $\|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{i0}\|_1 > \eta$ which implies that $\hat{\boldsymbol{\theta}}_i \notin \mathcal{B}_i(\eta)$ for all $1 \leq i \leq N$. It follows that $Q_{i,T}(\hat{\boldsymbol{\theta}}_i) \leq Q_{i,T}(\boldsymbol{\theta}_{i0})$ for some $1 \leq i \leq N$ by definition of $\hat{\boldsymbol{\theta}}_i = \arg\min\{Q_{i,T}(\boldsymbol{\theta}_i)\}$, which is equivalent to (2.18).

Note that $\hat{\boldsymbol{\theta}}_i \notin \mathcal{B}_i(\eta)$ implies $Q_{i,T}(\hat{\boldsymbol{\theta}}_i) \leq 0$ by definition. Thus, by equation (S.1.5), the following inclusion relationships are true:

$$\left\{\max_{1\leq i\leq N} \|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{i0}\|_1 > \eta\right\} \subseteq \{Q_{i,T}(\boldsymbol{\theta}_i) \leq 0, \exists \boldsymbol{\theta}_i \notin \mathcal{B}_i(\eta)\} \subseteq \bigcup_{i=1}^{N}\left\{\sup_{\boldsymbol{\theta}_i\in\mathcal{B}_i(\eta)} |\Delta_i(\boldsymbol{\theta}_i) - E(\Delta_i(\boldsymbol{\theta}_i))| \geq \epsilon_\eta\right\}.$$

because $Q_{i,T}(\boldsymbol{\theta}_{i0}) = 0$. It follows that,

$$P\left\{\max_{1\leq i\leq N} \|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{i0}\|_1 > \eta\right\} \leq N \max_{1\leq i\leq N} P\left\{\sup_{\boldsymbol{\theta}_i\in\mathcal{B}_i(\eta)} |\Delta_i(\boldsymbol{\theta}_i) - E(\Delta_i(\boldsymbol{\theta}_i))| \geq \epsilon_\eta\right\}.$$

We therefore need to show that

$$\max_{1\leq i\leq N} P\left\{\sup_{\boldsymbol{\theta}_i\in\mathcal{B}_i(\eta)} |\Delta_i(\boldsymbol{\theta}_i) - E(\Delta_i(\boldsymbol{\theta}_i))| \geq \epsilon_\eta\right\} = o(N^{-1}), \tag{S.1.6}$$

which is similar to equation (A.3) in Kato, Galvao and Montes-Rojas (2012) and equation (15) in Galvao and Wang (2015). Recall that as $N \to \infty$, automatically $T \to \infty$ too.

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

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

$$|\Delta(\boldsymbol{\theta}) - E(\Delta(\boldsymbol{\theta}))| \leq |\Delta(\boldsymbol{\theta}^{(j)}) - E(\Delta(\boldsymbol{\theta}^{(j)}))| + \frac{2\epsilon}{3}. \tag{S.1.7}$$

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

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

Therefore, following (S.1.7), we write,

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

By Hoeffding's inequality, each probability can be bounded by $2\exp\left(-(\epsilon/3)^2(T/2M^2)\right)$, and therefore,

$$P\left(\sup_{\boldsymbol{\theta} \in \mathcal{B}(\eta)} |\Delta_i(\boldsymbol{\theta}) - E(\Delta_i(\boldsymbol{\theta}))| > \epsilon\right) \leq 2K\exp(-DT) = O(\exp(-T)), \tag{S.1.8}$$

where D is a constant that depends on ϵ and not on i. If $\log(N)/T \to 0$, then $O(\exp(-T)) = o(N^{-1})$, which completes the proof. □

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

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

The first term converges in probability to zero as established in Theorem 1 and the last equality follows by Assumption 5. □

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

$$\begin{aligned}\hat{\boldsymbol{\vartheta}}(\tau) - \boldsymbol{\vartheta}(\tau) &= \frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{\vartheta}}_i(\tau) - \boldsymbol{\vartheta}(\tau)) = \frac{1}{N}\sum_{i=1}^{N}((\hat{\boldsymbol{\vartheta}}_i(\tau) - \boldsymbol{\vartheta}_i(\tau)) + (\boldsymbol{\vartheta}_i(\tau) - \boldsymbol{\vartheta}(\tau))) \\ &= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Xi}_i \circ [(\hat{\boldsymbol{\pi}}_i(\tau) - \boldsymbol{\pi}_i(\tau)) + (\boldsymbol{\pi}_i(\tau) - \boldsymbol{\pi}(\tau))] \qquad \text{(S.1.9)}\end{aligned}$$

It follows that,

$$\begin{aligned}\sqrt{N}\left(\hat{\boldsymbol{\vartheta}}(\tau) - \boldsymbol{\vartheta}(\tau)\right) &= \frac{\sqrt{N}}{N}\sum_{i=1}^{N}\boldsymbol{\Xi}_i \circ (\hat{\boldsymbol{\pi}}_i(\tau) - \boldsymbol{\pi}_i(\tau)) + \frac{\sqrt{N}}{N}\sum_{i=1}^{N}\boldsymbol{\Xi}_i \circ (\boldsymbol{\pi}_i(\tau) - \boldsymbol{\pi}(\tau)) \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\theta}}_{1i}(\tau) - \boldsymbol{\theta}_{1i}(\tau)\right) + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\boldsymbol{\theta}_{1i}(\tau) - \boldsymbol{\theta}_1(\tau)), \qquad \text{(S.1.10)}\end{aligned}$$

where $\boldsymbol{\theta}_{1i}(\tau) = (\lambda_i(\tau), \boldsymbol{\beta}_i(\tau)')'$ denote the $(1+p_x)$ first elements in $\boldsymbol{\theta}_i(\tau) = (\boldsymbol{\theta}_{1i}(\tau)', \boldsymbol{\theta}_{2i}(\tau)')'$, and by definition, $\boldsymbol{\theta}_{1i}(\tau) = \boldsymbol{\Xi}_i \circ \boldsymbol{\pi}_i(\tau)$. By Theorem 1, $\hat{\boldsymbol{\theta}}_{1i}(\tau) - \boldsymbol{\theta}_{1i}(\tau) = o_p(1)$.

We now obtain the asymptotic representation of $\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_i(\tau)$ following closely Galvao and Wang (2015). Define

$$\mathbb{H}_i(\boldsymbol{\theta}_i) = \frac{1}{T}\sum_{t=1}^{T}\psi_\tau(y_{it} - \mathbf{W}_{it}'\boldsymbol{\theta}_i)\mathbf{W}_{it}$$

and $H_i(\boldsymbol{\theta}_i) = E(\mathbb{H}_i(\boldsymbol{\theta}_i))$. We use an expansion of $H_i(\hat{\boldsymbol{\theta}}_i)$ around $\boldsymbol{\theta}_{i0}$ to obtain,

$$H_i(\hat{\boldsymbol{\theta}}_i) = H_i(\boldsymbol{\theta}_{i0}) + \mathbf{J}_i(\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau)) + O_p\left[(\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau))^2\right],$$

where $\mathbf{J}_i := \partial H_i(\boldsymbol{\theta}_i)/\partial \boldsymbol{\theta}_{i0} = E(g_i(0|\mathbf{W}_{it})\mathbf{W}_{it}\mathbf{W}'_{it}))$. Basic manipulations lead to:

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

For fixed N, the second term in the last expression is $o_p(1)$. In the case of panel data, we need to find the order of

$$\max_{1\leq i\leq N}\left[(\mathbb{H}_i(\hat{\boldsymbol{\theta}}_i) - \mathbb{H}_i(\boldsymbol{\theta}_{i0})) - (H_i(\hat{\boldsymbol{\theta}}_i) - H_i(\boldsymbol{\theta}_{i0}))\right].$$

Lemma S.1 establishes that order. Moreover, by the computational property of quantile regression, $\mathbb{H}_i(\hat{\boldsymbol{\theta}}_i(\tau)) = O_p(T^{-1})$. Therefore, for each $1 \leq i \leq N$, we have

$$\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau) = -\mathbf{J}_i^{-1}\mathbb{H}_i(\boldsymbol{\theta}_{i0}) + O_p(d_N) + O_p(T^{-1}) + \mathbf{J}_i^{-1}O_p\left((\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau))^2\right), \quad \text{(S.1.11)}$$

After basic simplifications, we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\left[\sqrt{N}\left(\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau)\right)\right] = -\frac{1}{N}\sum_{i=1}^{N}\mathbf{J}_i^{-1}\sqrt{N}\mathbb{H}_i(\boldsymbol{\theta}_i) + \sqrt{N}O_p(d_N). \quad \text{(S.1.12)}$$

The first term is $O_p(T^{-1/2})$ and second term is $O_p(N^{1/2}d_N)$. Using Lemma S.1, we have that $N^{1/2}d_N = N^{1/2}\log(N)^{3/4}T^{-3/4}$. Therefore, if $N^{2/3}\log(N)/T \to 0$, the second term is asymptotically negligible. This implies too that the first term in (S.1.10) is asymptotically negligible as N and $T \to \infty$.

Therefore, by standard arguments, as N and T tends to infinity under the conditions of Theorem 1, $\sqrt{N}\left(\hat{\boldsymbol{\vartheta}}(\tau) - \boldsymbol{\vartheta}(\tau)\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{V}_v)$. □

Proof of Theorem 4. If $\boldsymbol{\vartheta}_i(\tau) = \boldsymbol{\vartheta}(\tau)$ for $1 \leq i \leq N$, equation (S.1.9) can be written as

$$\hat{\boldsymbol{\vartheta}}(\tau) - \boldsymbol{\vartheta}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Xi}_i \circ (\hat{\boldsymbol{\pi}}_i(\tau) - \boldsymbol{\pi}(\tau)). \quad \text{(S.1.13)}$$

Using the definitions introduced in the proof of Theorem 3, we write

$$\begin{aligned}\sqrt{NT}(\hat{\boldsymbol{\vartheta}}(\tau)-\boldsymbol{\vartheta}(\tau)) &= \frac{\sqrt{NT}}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{\theta}}_{1i}(\tau)-\boldsymbol{\theta}_1(\tau)) \\ &= \sqrt{NT}(\hat{\boldsymbol{\theta}}_1(\tau)-\boldsymbol{\theta}_1(\tau)) \qquad \text{(S.1.14)}\end{aligned}$$

where $\hat{\boldsymbol{\theta}}_1(\tau) = N^{-1}\sum_{i=1}^{N}\hat{\boldsymbol{\theta}}_{1i}(\tau)$. Following Theorem 3 and Lemma S.1, we have, for each $1 \le i \le N$,

$$\hat{\boldsymbol{\theta}}_i(\tau)-\boldsymbol{\theta}_{i0}(\tau) = -\mathbf{J}_i^{-1}\mathbb{H}_i(\boldsymbol{\theta}_{i0}) + O_p(d_N) + O_p(T^{-1}) + \mathbf{J}_i^{-1}O_p\left((\hat{\boldsymbol{\theta}}_i(\tau)-\boldsymbol{\theta}_{i0}(\tau))^2\right), \quad \text{(S.1.15)}$$

because $\boldsymbol{\vartheta}_i(\tau) = \boldsymbol{\vartheta}(\tau)$ implies $\boldsymbol{\theta}_i(\tau) = (\boldsymbol{\theta}_1(\tau)', \boldsymbol{\theta}_{2i}(\tau)')'$ for $1 \le i \le N$. Again, using Lemma S.1, after basic simplifications, we obtain

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right) = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{J}_i^{-1}\psi_\tau(y_{it}-\mathbf{W}_{it}'\boldsymbol{\theta}_0)\mathbf{W}_{it} + O_p((T/\log(N))^{-3/4}).$$

Therefore, if $N^2(\log(N))^3/T \to 0$, $\|\hat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\| = O_p((NT)^{-1/2})$.

As N and T tends to infinity under the conditions of Theorem 1,

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right) = \sqrt{NT}\begin{pmatrix}\hat{\boldsymbol{\theta}}_1(\tau)-\boldsymbol{\theta}_1(\tau)\\ \hat{\boldsymbol{\theta}}_2(\tau)-\boldsymbol{\theta}_2(\tau)\end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0},\mathbf{V}), \qquad \text{(S.1.16)}$$

where $\mathbf{V}$ is the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}(\tau)$. Using equation (S.1.14), we conclude that $\sqrt{NT}(\hat{\boldsymbol{\vartheta}}(\tau)-\boldsymbol{\vartheta}(\tau)) \xrightarrow{d} \mathcal{N}(\mathbf{0},\mathbf{V}_\psi)$, where $\mathbf{V}_\psi$ is the upper diagonal $(p_x+1)\times(p_x+1)$ block of $\mathbf{V}$. □

S.1.4. **Additional mathematical results**

Proposition S.1. *Let* $\mathbf{S}_{iT} = T^{-1}\sum_{t=1}^{T}\mathbf{X}_{it}\mathbf{X}_{it}'$ *and assume that there exists* T_0 *such that for all* $T > T_0$, $\inf_i \zeta_{\min}(\mathbf{S}_{iT}) > 0$, *and* $\sup_i \zeta_{\min}(\mathbf{S}_{iT}) > K$. *As* $(N, T, p_T) \to \infty$ *with* $p_T^3/T \to \varkappa$, $0 < \varkappa < \infty$, $\mathbf{S}_{iT} \xrightarrow{p} \mathbf{S}_i = E(\mathbf{X}_{it}\mathbf{X}_{it}')$ *such that* $\inf_i(\zeta_{\min}(\mathbf{S}_i)) > 0$ *for all* $1 \le i \le N$.

Proof. The proof is implicit in the proof of Theorem 1 in Chudik and Pesaran (2015), and therefore we refer the reader to equation (A.71) on page 418. See also footnote 11 in Chudik and Pesaran (2015). □

Lemma S.1. *Under Assumption 1 and Assumptions S.1-S.3, for* δ_N *such that* $\lim_{N\to\infty}\delta_N = 0$, *we have that*

$$\max_{1\leq i\leq N}\{[\mathbb{H}_i(\hat{\boldsymbol{\theta}}_i) - \mathbb{H}_i(\boldsymbol{\theta}_{i0})] - [H_i(\hat{\boldsymbol{\theta}}_i) - H_i(\boldsymbol{\theta}_{i0})]\} = O_p(d_N),$$

where $d_N = \log(N)^{1/4}T^{-3/4}\sqrt{|\log(\delta_N)|}$, *and*

$$\max_{1\leq i\leq N}\|\hat{\boldsymbol{\theta}}_i(\tau) - \boldsymbol{\theta}_{i0}(\tau)\| = O_p(\sqrt{\log(N)/T}).$$

Proof. The first result is obtained following Lemma 4 in Galvao and Wang (2015). The second result follows directly from Lemma 5 in Galvao and Wang (2015). □

S.2. **Monte Carlo**

This section reports results of several additional simulation exercises on the small sample performance of the proposed estimator, complementing the results reported in Section 3 of the paper. Observations on y_{it} for $i = 1, 2, \ldots, N$ and $t = -S+1, -S+2, .., 0, 1, ..., T$ are generated according to the model with two factors considered in Section 3.

As in Section 3, we assume that the error term u_{it} is an i.i.d. random variable distributed as Standard Normal. We expand the evidence by also considering that u_{it} is an i.i.d. random variable distributed as t-student with 4 degrees of freedom (t_4), and as χ^2 with 3 degrees of freedom (χ^2_3). We consider the following four variations of the model (with $\lambda_i = \lambda$):

Design 1: (Location shift model with homogeneous slopes). We consider $\beta_1 = 1$ in a location shift model with $\kappa_{1i} = 0$ for all $1 \leq i \leq N$.
Design 2: (Location shift model with heterogeneous slopes). We consider heterogeneous slope parameters $\beta_{1i} = \beta_1 + \nu_{1i}$ in a location shift model, where $\kappa_{1i} = 0$ for all $1 \leq i \leq N$, $\beta_1 = 1$ and $\nu_{1i} \sim \mathcal{U}(-0.25, 0.25)$. The parameter $\beta_{1i}(\tau) = \beta_{1i}$ for all i and τ.
Design 3: (Location-scale shift model with homogeneous slopes). We consider homogenous slope parameters $\beta_1 = 1$ in a location-scale shift model with $\kappa_{1i} \sim \mathcal{U}(0, 0.2)$. In this case, the slope parameter $\beta_{1i}(\tau) = \beta_1 + \kappa_{0i}\kappa_{1i}F_u^{-1}(\tau)$ and $E(\beta_{1i}(\tau)) = \beta_1 + 0.1F_u^{-1}(\tau)$.
Design 4: (Location-scale shift model with heterogeneous slopes). We consider heterogeneous slope parameters as in Design 2, $\beta_{1i} = \beta_1 + \nu_{1i}$, in a location-scale shift model with $\kappa_{1i} \sim \mathcal{U}(0, 0.2)$. We assume $\beta_1 = 1$ and $\nu_{1i} \sim \mathcal{U}(-0.25, 0.25)$ which implies that $\beta_{1i}(\tau) = \beta_{1i} + \kappa_{0i}\kappa_{1i}F_u^{-1}(\tau) = 1 + \nu_{1i} + \kappa_{0i}\kappa_{1i}F_u^{-1}(\tau)$. In this case, $E(\beta_{1i}(\tau)) = \beta_1(\tau) = 1 + 0.1F_u^{-1}(\tau)$.

Tables S.1 to Table S.2 present the bias and root mean square error (RMSE) for the slope parameter $\beta_1(\tau)$ in the location shift model with $\lambda = 0.5$. The summary results for other choices of λ are available upon request. While Table S.1 presents results for Designs 1 and 2, Table S.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 QMG estimator with the instrumental variable quantile regression estimator for dynamic panel data model developed by Galvao (2011), using $y_{i,t-2}$ as an instrument for $y_{i,t-1}$. This estimator is denoted by DQR. The QMG, is computed as the simple cross sectional average of standard quantile estimators, $\hat{\beta}_{1i}(\tau)$, using $\bar{\mathbf{z}}_t = (\bar{y}_t, \bar{y}_{t-1}, \bar{\mathbf{x}}_t')'$ to proxy the true unobserved factors f_{1t} and f_{2t}. We do not consider other existing quantile estimators, such as the classical quantile regression estimator, the fixed effects minimum distance quantile regression estimator by Galvao and Wang (2015), and the penalized quantile regression estimator, since all these estimators are biased when the model includes a lagged dependent variable. Therefore, we restrict our comparison to DQR, which is the only estimator in the literature proposed for dynamic panel quantile regression models.

			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
N	T		Design 1: Location shift with homogeneous slopes											
100	50	Bias	-0.191	0.053	-0.187	0.055	-0.171	0.057	-0.166	0.063	-0.109	0.080	-0.109	0.046
100	50	RMSE	0.221	0.061	0.218	0.064	0.200	0.065	0.196	0.073	0.152	0.097	0.148	0.063
100	100	Bias	-0.253	0.022	-0.249	0.023	-0.230	0.027	-0.230	0.032	-0.173	0.040	-0.164	0.018
100	100	RMSE	0.270	0.029	0.266	0.032	0.247	0.034	0.247	0.040	0.194	0.053	0.184	0.031
100	200	Bias	-0.293	0.003	-0.291	0.003	-0.277	0.005	-0.275	0.007	-0.217	0.020	-0.204	0.005
100	200	RMSE	0.303	0.015	0.301	0.016	0.287	0.015	0.285	0.018	0.229	0.031	0.216	0.018
200	50	Bias	-0.198	0.056	-0.195	0.057	-0.176	0.059	-0.170	0.066	-0.116	0.085	-0.115	0.048
200	50	RMSE	0.225	0.060	0.221	0.061	0.203	0.063	0.197	0.072	0.153	0.094	0.149	0.057
200	100	Bias	-0.271	0.028	-0.267	0.028	-0.235	0.031	-0.233	0.035	-0.174	0.047	-0.167	0.023
200	100	RMSE	0.286	0.031	0.283	0.031	0.252	0.034	0.250	0.038	0.192	0.053	0.185	0.029
200	200	Bias	-0.294	0.011	-0.292	0.011	-0.281	0.011	-0.280	0.014	-0.208	0.023	-0.198	0.010
200	200	RMSE	0.303	0.015	0.301	0.015	0.290	0.015	0.289	0.018	0.219	0.028	0.208	0.015
N	T		Design 2: Location shift with heterogeneous slopes											
100	50	Bias	-0.195	0.049	-0.190	0.049	-0.170	0.054	-0.165	0.061	-0.114	0.081	-0.115	0.045
100	50	RMSE	0.224	0.058	0.220	0.060	0.206	0.065	0.201	0.074	0.158	0.099	0.156	0.063
100	100	Bias	-0.267	0.021	-0.265	0.022	-0.237	0.024	-0.235	0.028	-0.179	0.042	-0.171	0.020
100	100	RMSE	0.283	0.033	0.280	0.035	0.255	0.035	0.253	0.040	0.199	0.058	0.190	0.036
100	200	Bias	-0.296	0.001	-0.294	0.001	-0.279	0.006	-0.280	0.008	-0.212	0.021	-0.201	0.007
100	200	RMSE	0.307	0.021	0.305	0.022	0.290	0.021	0.290	0.023	0.226	0.034	0.215	0.022
200	50	Bias	-0.200	0.055	-0.196	0.054	-0.175	0.058	-0.169	0.065	-0.114	0.082	-0.110	0.047
200	50	RMSE	0.226	0.060	0.223	0.060	0.199	0.064	0.194	0.071	0.148	0.093	0.142	0.058
200	100	Bias	-0.258	0.028	-0.255	0.029	-0.245	0.029	-0.242	0.032	-0.181	0.046	-0.174	0.024
200	100	RMSE	0.272	0.033	0.270	0.034	0.260	0.033	0.257	0.038	0.200	0.052	0.192	0.030
200	200	Bias	-0.301	0.011	-0.299	0.010	-0.277	0.012	-0.277	0.014	-0.214	0.021	-0.203	0.009
200	200	RMSE	0.309	0.018	0.307	0.018	0.285	0.019	0.285	0.021	0.225	0.028	0.214	0.017

TABLE S.1. *Bias and root mean square error (RMSE) of quantile regression estimators for* $\beta_1(\tau)$ *in Designs 1 and 2. In all the variations of the model,* $\lambda = 0.5$*. DQR denotes the instrumental variable quantile regression estimator for dynamic quantile regression, and QMG denotes the proposed mean quantile group estimator.*

			Normal Distribution				t_4 distribution				χ^2_3 distribution			
			$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.25$	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
N	T		Design 3: Location-scale shift with homogeneous slopes											
100	50	Bias	-0.188	0.055	-0.165	0.070	-0.168	0.057	-0.145	0.077	-0.131	0.098	-0.102	0.075
100	50	RMSE	0.219	0.063	0.198	0.078	0.197	0.067	0.179	0.086	0.178	0.116	0.150	0.089
100	100	Bias	-0.250	0.023	-0.225	0.032	-0.226	0.027	-0.207	0.042	-0.203	0.044	-0.163	0.030
100	100	RMSE	0.267	0.031	0.244	0.039	0.243	0.035	0.226	0.049	0.227	0.059	0.187	0.040
100	200	Bias	-0.291	0.003	-0.268	0.009	-0.272	0.005	-0.251	0.014	-0.251	0.020	-0.205	0.010
100	200	RMSE	0.300	0.015	0.278	0.019	0.282	0.015	0.262	0.022	0.265	0.034	0.220	0.022
200	50	Bias	-0.195	0.057	-0.172	0.071	-0.171	0.060	-0.148	0.081	-0.139	0.101	-0.107	0.076
200	50	RMSE	0.222	0.061	0.202	0.075	0.199	0.065	0.178	0.086	0.177	0.110	0.147	0.082
200	100	Bias	-0.268	0.029	-0.243	0.036	-0.231	0.031	-0.210	0.043	-0.204	0.053	-0.165	0.034
200	100	RMSE	0.283	0.032	0.260	0.038	0.247	0.034	0.228	0.046	0.224	0.059	0.187	0.040
200	200	Bias	-0.291	0.011	-0.268	0.016	-0.276	0.012	-0.256	0.020	-0.242	0.025	-0.199	0.015
200	200	RMSE	0.300	0.015	0.277	0.019	0.285	0.015	0.265	0.023	0.254	0.032	0.212	0.020
N	T		Design 4: Location-scale shift with heterogeneous slopes											
100	50	Bias	-0.194	0.052	-0.167	0.064	-0.166	0.055	-0.143	0.077	-0.139	0.096	-0.108	0.073
100	50	RMSE	0.223	0.061	0.201	0.074	0.202	0.067	0.184	0.088	0.187	0.114	0.159	0.087
100	100	Bias	-0.264	0.022	-0.241	0.030	-0.232	0.025	-0.211	0.039	-0.210	0.047	-0.171	0.031
100	100	RMSE	0.280	0.034	0.258	0.042	0.250	0.036	0.231	0.048	0.231	0.064	0.194	0.045
100	200	Bias	-0.292	0.002	-0.270	0.007	-0.275	0.006	-0.256	0.015	-0.246	0.020	-0.202	0.012
100	200	RMSE	0.304	0.022	0.282	0.023	0.285	0.021	0.267	0.027	0.262	0.036	0.219	0.026
200	50	Bias	-0.198	0.057	-0.174	0.068	-0.170	0.060	-0.147	0.079	-0.136	0.098	-0.103	0.074
200	50	RMSE	0.225	0.062	0.203	0.073	0.195	0.065	0.173	0.085	0.173	0.110	0.142	0.083
200	100	Bias	-0.254	0.029	-0.231	0.036	-0.241	0.029	-0.219	0.040	-0.212	0.050	-0.172	0.034
200	100	RMSE	0.269	0.034	0.247	0.041	0.257	0.034	0.236	0.045	0.233	0.057	0.194	0.039
200	200	Bias	-0.298	0.011	-0.274	0.015	-0.272	0.013	-0.253	0.019	-0.249	0.022	-0.204	0.014
200	200	RMSE	0.306	0.019	0.283	0.022	0.281	0.019	0.262	0.025	0.261	0.031	0.217	0.021

TABLE S.2. *Bias and root mean square error (RMSE) of quantile regression estimators for $\beta_1(\tau)$ in Designs 3 and 4. In all the variations of the model, $\lambda = 0.5$. Also, see notes to Table S.1.*

S.2.1. **Bias and Root Mean Square Error**

Table S.1 shows that the DQR estimator of β_1 is biased. On the other hand, the performance of the QMG estimator is excellent, with biases in general lower than 10% for $T = 50$, and decreasing rapidly to 1% when $T = 200$. In all the variations of the model considered in the table, the QMG estimator performs much better than DQR in terms of RMSE, as well.

Table S.2 presents results for the location-scale shift model where $\beta_1(\tau)$ changes by quantile. We continue to see that the DQR estimator is biased and performs poorly in terms of RMSE. The performance of the QMG estimator in these variations of the model is similar to the results reported for the baseline model in Table S.1, with low biases and small RMSE. For values of T larger than 50, the bias of the proposed estimator is always negative and ranges between 0.7% and 4%, and its RMSE is substantially below that of the DQR estimator. The RMSE of QMG relative to DQR is around 30 percent for $N = 100, T = 50$, and falls to around 0.05 for $N = T = 200$. The relative efficiency of the QMG estimator is similar across all the four designs.

We expanded the simulation evidence for the slope parameter β_1 to consider different values of λ. In the online supplement we present results for $\lambda \in \{0.25, 0.75\}$ considering the same designs as in Tables S.1 and S.2, with $N = 100$ 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 in the heterogeneous case when λ_i is distributed as $\mathcal{U}[0.025, 0.925]$ below.

We now turn our attention to the estimates of $\lambda(\tau)$ and $\theta_1(\tau) = \beta_1(\tau)/(1 - \lambda(\tau))$. Tables S.3, S.4, S.5 and S.6 show the bias and RMSE of the DQR and QMG estimators for these parameters. 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 presents 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 panel presents results when $u_{it} \sim \chi^2_3$.

As before, the results indicate that the bias of the DQR estimator can be large, in particular for the long run coefficient θ_1. The QMG estimator offers nearly zero biases for large N

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ_1		Parameter: λ		Parameter: θ_1	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
N	T		Normal Distribution							
100	50	Bias	0.188	-0.059	0.644	-0.058	0.186	-0.059	0.639	-0.043
100	50	RMSE	0.199	0.061	0.702	0.085	0.197	0.062	0.700	0.080
100	100	Bias	0.221	-0.019	0.703	-0.011	0.219	-0.020	0.703	-0.007
100	100	RMSE	0.226	0.022	0.736	0.044	0.224	0.023	0.737	0.046
100	200	Bias	0.240	-0.002	0.738	0.007	0.239	-0.002	0.736	0.008
100	200	RMSE	0.243	0.008	0.758	0.029	0.242	0.009	0.758	0.030
200	50	Bias	0.194	-0.063	0.666	-0.069	0.192	-0.063	0.665	-0.054
200	50	RMSE	0.204	0.064	0.722	0.083	0.202	0.064	0.724	0.071
200	100	Bias	0.231	-0.026	0.734	-0.027	0.229	-0.026	0.731	-0.023
200	100	RMSE	0.235	0.027	0.760	0.038	0.234	0.027	0.759	0.036
200	200	Bias	0.242	-0.009	0.744	-0.006	0.241	-0.009	0.740	-0.004
200	200	RMSE	0.244	0.010	0.757	0.017	0.243	0.010	0.754	0.018
N	T		t_4 distribution							
100	50	Bias	0.177	-0.062	0.607	-0.066	0.175	-0.070	0.614	-0.060
100	50	RMSE	0.187	0.064	0.677	0.092	0.186	0.072	0.684	0.096
100	100	Bias	0.208	-0.023	0.664	-0.019	0.209	-0.027	0.669	-0.018
100	100	RMSE	0.214	0.025	0.698	0.044	0.214	0.029	0.704	0.049
100	200	Bias	0.231	-0.004	0.699	0.002	0.230	-0.006	0.702	0.000
100	200	RMSE	0.233	0.008	0.723	0.028	0.233	0.010	0.725	0.030
200	50	Bias	0.177	-0.067	0.593	-0.083	0.176	-0.075	0.599	-0.073
200	50	RMSE	0.187	0.068	0.654	0.093	0.186	0.075	0.659	0.107
200	100	Bias	0.209	-0.027	0.651	-0.028	0.209	-0.031	0.658	-0.027
200	100	RMSE	0.214	0.028	0.683	0.038	0.214	0.032	0.690	0.041
200	200	Bias	0.232	-0.009	0.695	-0.007	0.232	-0.011	0.701	-0.008
200	200	RMSE	0.234	0.010	0.710	0.018	0.235	0.012	0.717	0.022
N	T		χ^2_3 distribution							
100	50	Bias	0.135	-0.091	0.473	-0.102	0.134	-0.051	0.466	-0.058
100	50	RMSE	0.149	0.092	0.529	0.146	0.147	0.053	0.519	0.104
100	100	Bias	0.172	-0.038	0.543	-0.043	0.165	-0.018	0.515	-0.018
100	100	RMSE	0.179	0.040	0.579	0.080	0.171	0.019	0.551	0.052
100	200	Bias	0.195	-0.016	0.576	-0.011	0.185	-0.005	0.540	-0.003
100	200	RMSE	0.198	0.017	0.597	0.046	0.189	0.007	0.560	0.033
200	50	Bias	0.140	-0.092	0.480	-0.098	0.136	-0.052	0.455	-0.056
200	50	RMSE	0.152	0.093	0.522	0.124	0.147	0.053	0.492	0.079
200	100	Bias	0.173	-0.043	0.542	-0.044	0.166	-0.021	0.511	-0.020
200	100	RMSE	0.179	0.043	0.565	0.063	0.172	0.021	0.533	0.040
200	200	Bias	0.191	-0.019	0.576	-0.018	0.182	-0.008	0.534	-0.006
200	200	RMSE	0.194	0.020	0.590	0.036	0.185	0.009	0.549	0.023

TABLE S.3. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ_1 in Design 1. In all the variations of the model, $\lambda = 0.5$. Also, see notes to Table S.1.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ_1		Parameter: λ		Parameter: θ_1	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
N	T		Normal Distribution							
100	50	Bias	0.193	-0.059	0.668	-0.063	0.191	-0.057	0.670	-0.038
100	50	RMSE	0.203	0.061	0.735	0.091	0.201	0.059	0.739	0.110
100	100	Bias	0.229	-0.020	0.732	-0.014	0.228	-0.020	0.729	-0.008
100	100	RMSE	0.234	0.023	0.770	0.054	0.232	0.023	0.770	0.055
100	200	Bias	0.242	-0.002	0.740	0.005	0.241	-0.002	0.741	0.006
100	200	RMSE	0.244	0.008	0.759	0.040	0.244	0.009	0.760	0.042
200	50	Bias	0.194	-0.064	0.674	-0.072	0.193	-0.063	0.678	-0.058
200	50	RMSE	0.205	0.065	0.741	0.087	0.204	0.064	0.752	0.078
200	100	Bias	0.223	-0.026	0.711	-0.025	0.222	-0.026	0.712	-0.020
200	100	RMSE	0.228	0.027	0.744	0.040	0.227	0.027	0.748	0.039
200	200	Bias	0.244	-0.009	0.742	-0.006	0.243	-0.009	0.739	-0.006
200	200	RMSE	0.246	0.010	0.758	0.028	0.245	0.010	0.756	0.028
N	T		t_4 distribution							
100	50	Bias	0.174	-0.062	0.591	-0.070	0.173	-0.069	0.600	-0.061
100	50	RMSE	0.186	0.064	0.661	0.102	0.185	0.072	0.669	0.106
100	100	Bias	0.211	-0.022	0.669	-0.019	0.210	-0.026	0.667	-0.019
100	100	RMSE	0.217	0.024	0.702	0.053	0.216	0.028	0.703	0.059
100	200	Bias	0.231	-0.004	0.696	0.002	0.232	-0.006	0.700	0.003
100	200	RMSE	0.234	0.009	0.716	0.041	0.235	0.011	0.722	0.042
200	50	Bias	0.178	-0.066	0.598	-0.078	0.176	-0.074	0.597	-0.072
200	50	RMSE	0.187	0.067	0.642	0.092	0.185	0.075	0.645	0.091
200	100	Bias	0.215	-0.027	0.678	-0.030	0.215	-0.031	0.685	-0.030
200	100	RMSE	0.220	0.028	0.708	0.045	0.220	0.032	0.716	0.047
200	200	Bias	0.232	-0.010	0.705	-0.006	0.232	-0.012	0.706	-0.009
200	200	RMSE	0.234	0.011	0.719	0.029	0.234	0.013	0.722	0.031
N	T		χ^2_3 distribution							
100	50	Bias	0.139	-0.090	0.486	-0.096	0.137	-0.051	0.463	-0.056
100	50	RMSE	0.152	0.091	0.540	0.142	0.149	0.052	0.513	0.100
100	100	Bias	0.178	-0.039	0.568	-0.041	0.171	-0.018	0.535	-0.017
100	100	RMSE	0.184	0.040	0.601	0.083	0.176	0.020	0.567	0.060
100	200	Bias	0.193	-0.016	0.577	-0.009	0.185	-0.005	0.542	0.002
100	200	RMSE	0.196	0.017	0.600	0.052	0.188	0.007	0.564	0.042
200	50	Bias	0.138	-0.091	0.479	-0.099	0.135	-0.051	0.462	-0.056
200	50	RMSE	0.150	0.092	0.528	0.130	0.146	0.052	0.510	0.086
200	100	Bias	0.176	-0.042	0.546	-0.044	0.170	-0.020	0.517	-0.019
200	100	RMSE	0.182	0.042	0.571	0.065	0.175	0.021	0.541	0.042
200	200	Bias	0.195	-0.019	0.584	-0.020	0.186	-0.008	0.546	-0.007
200	200	RMSE	0.198	0.019	0.598	0.039	0.189	0.008	0.560	0.028

TABLE S.4. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ_1 in Design 2. In all the variations of the model, $\lambda = 0.5$. Also, see notes to Table S.1.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ_1		Parameter: λ		Parameter: θ_1	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
N	T		Normal Distribution							
100	50	Bias	0.187	-0.061	0.637	-0.060	0.183	-0.062	0.596	-0.015
100	50	RMSE	0.198	0.063	0.694	0.089	0.194	0.064	0.648	0.072
100	100	Bias	0.219	-0.020	0.695	-0.011	0.216	-0.021	0.649	0.009
100	100	RMSE	0.224	0.023	0.729	0.045	0.221	0.024	0.678	0.046
100	200	Bias	0.239	-0.002	0.729	0.007	0.237	-0.003	0.676	0.018
100	200	RMSE	0.241	0.009	0.749	0.029	0.240	0.009	0.695	0.035
200	50	Bias	0.192	-0.066	0.661	-0.073	0.189	-0.065	0.618	-0.024
200	50	RMSE	0.202	0.066	0.717	0.083	0.199	0.066	0.666	0.053
200	100	Bias	0.229	-0.027	0.724	-0.028	0.226	-0.027	0.673	-0.008
200	100	RMSE	0.233	0.028	0.750	0.039	0.231	0.028	0.696	0.030
200	200	Bias	0.240	-0.009	0.735	-0.006	0.238	-0.009	0.678	0.005
200	200	RMSE	0.242	0.011	0.748	0.018	0.240	0.011	0.690	0.018
N	T		t_4 distribution							
100	50	Bias	0.174	-0.063	0.595	-0.067	0.171	-0.071	0.556	-0.027
100	50	RMSE	0.184	0.065	0.664	0.094	0.182	0.073	0.614	0.078
100	100	Bias	0.206	-0.024	0.653	-0.020	0.204	-0.028	0.598	0.003
100	100	RMSE	0.211	0.025	0.686	0.046	0.209	0.030	0.628	0.048
100	200	Bias	0.228	-0.004	0.689	0.001	0.225	-0.007	0.619	0.013
100	200	RMSE	0.231	0.009	0.712	0.028	0.228	0.010	0.640	0.034
200	50	Bias	0.174	-0.069	0.583	-0.083	0.171	-0.076	0.544	-0.035
200	50	RMSE	0.185	0.069	0.640	0.094	0.181	0.077	0.594	0.069
200	100	Bias	0.206	-0.027	0.639	-0.028	0.204	-0.032	0.589	-0.008
200	100	RMSE	0.211	0.028	0.670	0.039	0.210	0.032	0.615	0.033
200	200	Bias	0.229	-0.009	0.684	-0.007	0.228	-0.012	0.619	0.005
200	200	RMSE	0.232	0.010	0.699	0.018	0.230	0.013	0.632	0.022
N	T		χ_3^2 distribution							
100	50	Bias	0.117	-0.084	0.433	-0.109	0.114	-0.048	0.413	-0.013
100	50	RMSE	0.130	0.085	0.488	0.161	0.126	0.049	0.457	0.093
100	100	Bias	0.150	-0.035	0.495	-0.051	0.141	-0.016	0.437	0.001
100	100	RMSE	0.156	0.036	0.531	0.093	0.147	0.018	0.466	0.053
100	200	Bias	0.170	-0.014	0.520	-0.016	0.159	-0.004	0.445	0.006
100	200	RMSE	0.173	0.015	0.541	0.055	0.162	0.007	0.465	0.036
200	50	Bias	0.120	-0.086	0.439	-0.108	0.115	-0.048	0.404	-0.015
200	50	RMSE	0.132	0.086	0.479	0.138	0.125	0.049	0.432	0.062
200	100	Bias	0.151	-0.039	0.496	-0.052	0.142	-0.019	0.437	-0.002
200	100	RMSE	0.156	0.039	0.520	0.073	0.148	0.019	0.455	0.038
200	200	Bias	0.167	-0.017	0.520	-0.022	0.156	-0.007	0.443	0.003
200	200	RMSE	0.170	0.018	0.535	0.045	0.159	0.008	0.457	0.025

TABLE S.5. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ_1 in Design 3. In all the variations of the model, $\lambda = 0.5$. Also, see notes to Table S.1.*

			$\tau = 0.50$ quantile				$\tau = 0.25$ quantile			
			Parameter: λ		Parameter: θ_1		Parameter: λ		Parameter: θ_1	
			DQR	QMG	DQR	QMG	DQR	QMG	DQR	QMG
N	T		Normal Distribution							
100	50	Bias	0.191	-0.061	0.662	-0.064	0.188	-0.059	0.631	-0.013
100	50	RMSE	0.202	0.063	0.729	0.094	0.199	0.062	0.692	0.084
100	100	Bias	0.227	-0.021	0.724	-0.016	0.225	-0.021	0.674	0.007
100	100	RMSE	0.232	0.024	0.763	0.056	0.230	0.024	0.710	0.055
100	200	Bias	0.240	-0.002	0.733	0.005	0.238	-0.002	0.679	0.015
100	200	RMSE	0.243	0.008	0.752	0.041	0.241	0.009	0.697	0.045
200	50	Bias	0.193	-0.066	0.664	-0.074	0.190	-0.065	0.635	-0.030
200	50	RMSE	0.203	0.067	0.731	0.090	0.201	0.066	0.695	0.061
200	100	Bias	0.221	-0.027	0.705	-0.026	0.219	-0.027	0.659	-0.005
200	100	RMSE	0.226	0.028	0.738	0.041	0.224	0.028	0.689	0.035
200	200	Bias	0.242	-0.010	0.734	-0.006	0.240	-0.009	0.677	0.003
200	200	RMSE	0.244	0.011	0.749	0.028	0.242	0.011	0.690	0.029
N	T		t_4 distribution							
100	50	Bias	0.171	-0.064	0.580	-0.071	0.169	-0.071	0.545	-0.029
100	50	RMSE	0.183	0.066	0.651	0.104	0.181	0.074	0.603	0.095
100	100	Bias	0.208	-0.023	0.657	-0.021	0.206	-0.027	0.598	0.002
100	100	RMSE	0.214	0.025	0.689	0.055	0.212	0.029	0.628	0.056
100	200	Bias	0.228	-0.005	0.685	0.001	0.228	-0.007	0.619	0.015
100	200	RMSE	0.231	0.009	0.705	0.041	0.230	0.011	0.639	0.045
200	50	Bias	0.175	-0.067	0.586	-0.079	0.171	-0.075	0.545	-0.039
200	50	RMSE	0.184	0.068	0.629	0.093	0.180	0.076	0.585	0.069
200	100	Bias	0.212	-0.027	0.664	-0.031	0.210	-0.031	0.609	-0.012
200	100	RMSE	0.217	0.028	0.693	0.046	0.216	0.032	0.634	0.038
200	200	Bias	0.229	-0.010	0.692	-0.006	0.227	-0.012	0.623	0.003
200	200	RMSE	0.231	0.011	0.707	0.029	0.229	0.013	0.638	0.030
N	T		χ^2_3 distribution							
100	50	Bias	0.121	-0.083	0.443	-0.107	0.116	-0.047	0.408	-0.015
100	50	RMSE	0.133	0.084	0.498	0.156	0.127	0.048	0.449	0.088
100	100	Bias	0.155	-0.036	0.517	-0.050	0.146	-0.017	0.451	0.002
100	100	RMSE	0.160	0.037	0.551	0.094	0.151	0.018	0.479	0.061
100	200	Bias	0.168	-0.014	0.521	-0.016	0.158	-0.004	0.449	0.010
100	200	RMSE	0.172	0.015	0.544	0.059	0.161	0.007	0.470	0.045
200	50	Bias	0.119	-0.084	0.439	-0.108	0.114	-0.047	0.409	-0.015
200	50	RMSE	0.130	0.085	0.486	0.143	0.124	0.048	0.443	0.072
200	100	Bias	0.153	-0.038	0.495	-0.053	0.145	-0.019	0.438	-0.002
200	100	RMSE	0.159	0.039	0.519	0.075	0.150	0.019	0.456	0.038
200	200	Bias	0.170	-0.017	0.529	-0.025	0.159	-0.007	0.453	0.001
200	200	RMSE	0.173	0.017	0.544	0.047	0.162	0.008	0.466	0.030

TABLE S.6. *Bias and root mean square error (RMSE) of quantile regression estimators for λ and θ_1 in Design 4. In all the variations of the model, $\lambda = 0.5$. Also, see notes to Table S.1.*

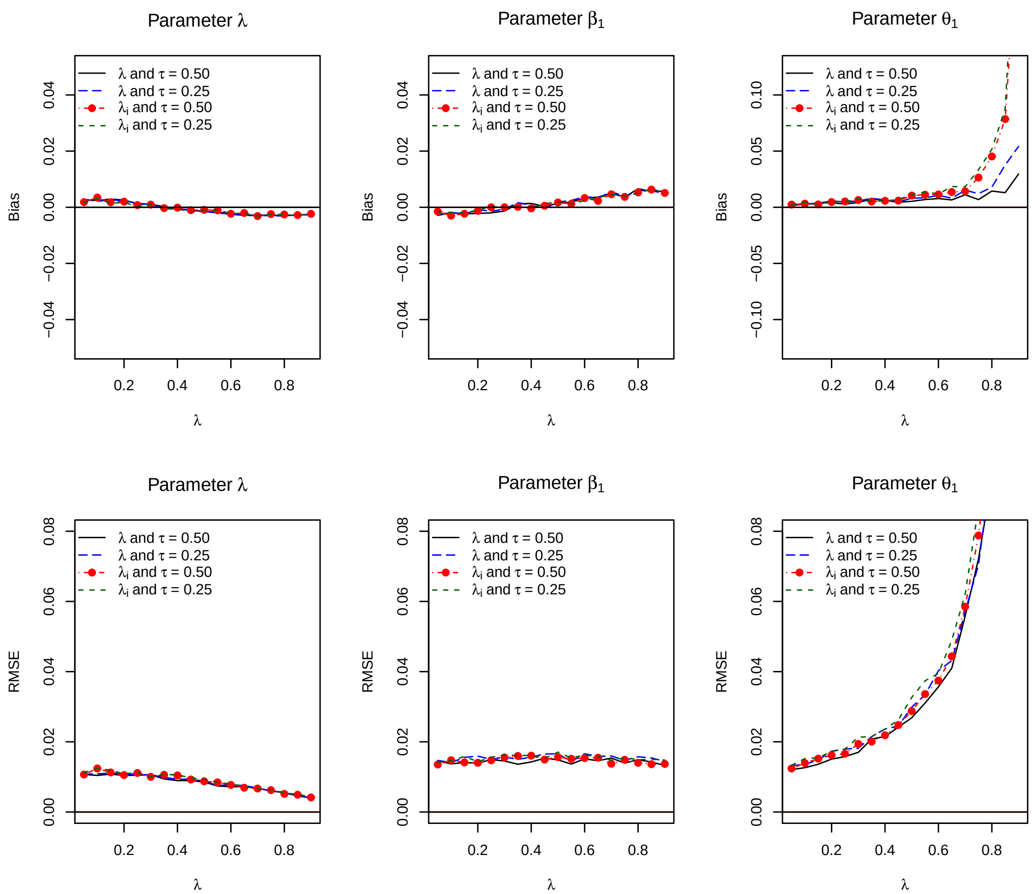

FIGURE S.1. *Small sample performance of the QMG estimator for different values of* λ. *The figure present Bias and RMSE of the QMG estimator for* $E(\lambda(\tau))$, $E(\beta_1(\tau))$ *and* $E(\theta_1(\tau))$ *at the 0.25 and 0.50 quantiles.*

and T. The DQR estimator is biased and its performance is not satisfactory in terms of both bias and RMSE. The location-scale shift case, presented in Tables S.5 and S.6, reveals similar findings.

Figure S.1 offers a visual display 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_1 and θ_1 for different true values of λ. We considered Design 1 with $N = 100$ and $T = 200$. Recall that when λ increases, θ_1 increases too. For instance, while $\lambda = 0$ gives $\theta_1 = \beta_1 = 1$, $\lambda = 0.9$ 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 β_1. The bias tends to increase slightly, but it is never larger than 1% for values of λ close to unity. We also find that the RMSE of the estimator of β_1 does not change with λ. On the other hand, we observe that the absolute value of the bias of the QMG estimator for θ_1 increases rapidly with $\lambda \to 1$. 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 to be expected.

Figure S.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_1 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 deteriorates for large values of $\lambda = E(\lambda_i)$. We see an increase in the variance of the estimator, but the bias for θ_1 remains, in absolute value, small for $E(\lambda_i) < 0.65$. As can be seen from Figure S.1, the parameter vector $(E(\lambda_i), \beta_1)$ can be estimated with small bias and excellent RMSE performance in the case of heterogeneous λ_i's, so long as N and T are sufficiently large, and $E(\lambda_i)$ is not too close to unity.

Finally, we investigate the relative performance of DQR and QMG in models with and without factor structure, i.e. $\sum_{j=1}^{2} \sigma_\gamma \gamma_{ji} f_{jt}$ in equation (3.1). As in Figure S.1, we generate data using Design 1 with $N = 100$ and $T = 200$. In contrast with the previous design, we generate $\gamma_{1i} \sim iid\mathcal{N}(0.5, 1)$ and $\gamma_{2i} \sim iid\mathcal{N}(0.5, 1)$, and we set σ_γ to take values in the interval $[0, 1]$. Naturally, when $\sigma_\gamma = 0$, the model does not include latent factors. Figure S.2 presents the bias and RMSE of the estimators for λ, β_1 and θ_1. Consistent again with expectations, when equation (3.1) does not include factors, the DQR estimator offers the best finite sample performance. However, as shown in the figure, the QMG performs reasonably well even when $\sigma_\gamma = 0$ and it offers the best performance in terms of bias and RMSE when the degree of parameter heterogeneity is not too small.

References

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.

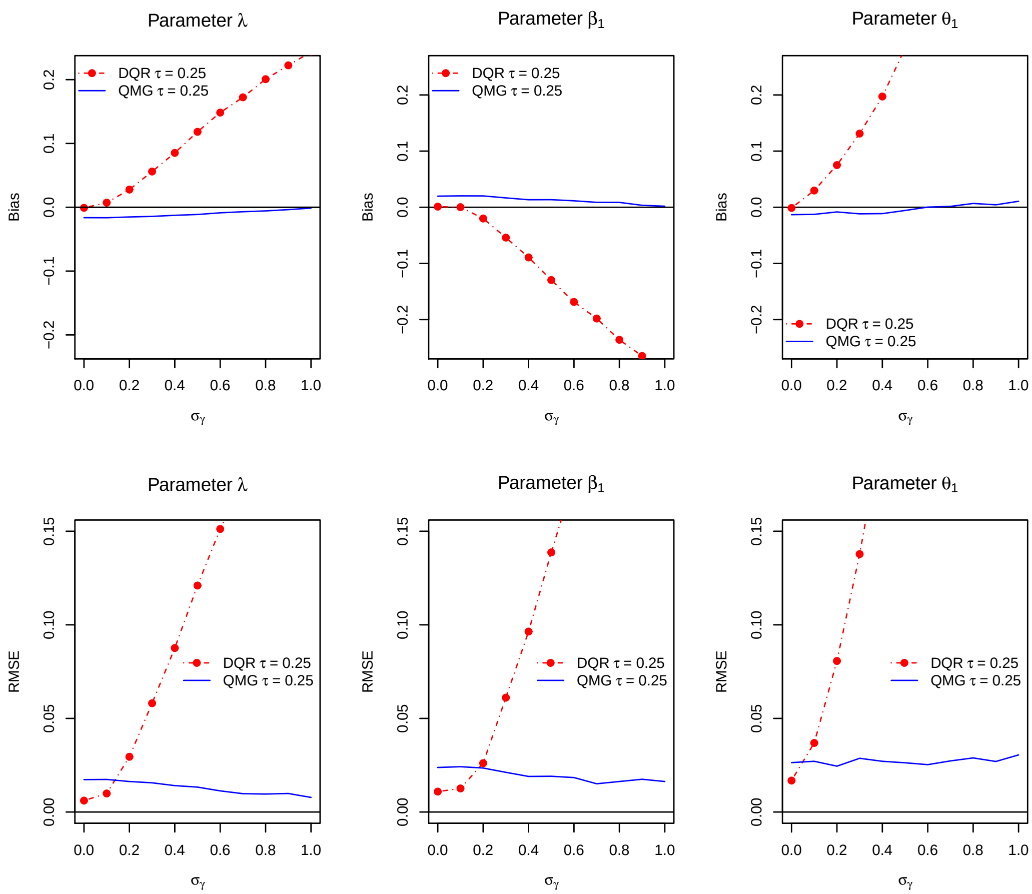

FIGURE S.2. *Small sample performance of the DQR and QMG estimators in models with and without latent factors.*

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

GALVAO, A. F., C. LAMARCHE, AND L. R. LIMA (2013): "Estimation of Censored Quantile Regression for Panel Data With Fixed Effects," *Journal of the American Statistical Association*, 108(503), 1075–1089.

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.

HE, X., AND Q.-M. SHAO (2000): "On Parameters of Increasing Dimensions," *Journal of Multivariate Analysis*, 73(1), 120–135.

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

Newey, W. K., and D. McFadden (1994): "Chapter 36 Large sample estimation and hypothesis testing," vol. 4 of *Handbook of Econometrics*, pp. 2111 – 2245. Elsevier.