

Graham, Bryan S.; Hahn, Jinyong; Poirier, Alexandre; Powell, James L.

Working Paper

Quantile regression with panel data

cemmap working paper, No. CWP12/15

Provided in Cooperation with:

The Institute for Fiscal Studies (IFS), London

Suggested Citation: Graham, Bryan S.; Hahn, Jinyong; Poirier, Alexandre; Powell, James L. (2015) : Quantile regression with panel data, cemmap working paper, No. CWP12/15, Centre for Microdata Methods and Practice (cemmap), London, <https://doi.org/10.1920/wp.cem.2015.1215>

This Version is available at:

<https://hdl.handle.net/10419/130026>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Supplemental Web Appendix to “Quantile regression with Panel Data” by Bryan S. Graham, Jinyong Hahn, Alexandre Poirier and James L. Powell: Proofs of Theorems 8, 9 and 10

This Supplemental Web Appendix contains proofs of Theorems 8, 9 and 10 that appear in Appendix A of the main paper. All notation is as established in, and equation numbering continues in sequence with that of, the main paper.

Proof of Theorem 8

Using arguments analogous to those used in Appendix B of the main paper we can show that $\hat{\delta}(\tau) - \delta(\tau)$ has the linear representation

$$\begin{aligned} \sqrt{N} \left(\hat{\delta}(\tau) - \delta(\tau) \right) &= \left(\frac{1}{N} \sum_{i=1}^N \mathbf{W}_i^{*'} \mathbf{W}_i^* \mathbf{1}(D_i = 0) \right)^{-1} \\ &\times \frac{1}{N} \sum_{i=1}^N \mathbf{W}_i^{*'} \mathbf{X}_i^* \mathbf{1}(D_i = 0) \sqrt{N} \left(\hat{Q}_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{X}_i) - Q_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{X}_i) \right) \\ &= \left(\sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{w}_l^* \hat{p}_l \right)^{-1} \sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{x}_l^* \hat{p}_l \sqrt{N} \left(\hat{Q}_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) - Q_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) \right) \end{aligned} \quad (103)$$

with

$$\sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{w}_l^* \hat{p}_l \xrightarrow{P} \sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{w}_l^* p_l = \mathbb{E}[\mathbf{W}^{*'} \mathbf{W}^* | D = 0] \pi_0$$

and

$$\sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{x}_l^* \hat{p}_l \sqrt{N} \left(\hat{Q}_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) - Q_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) \right) \xrightarrow{D} \sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{x}_l^* \sqrt{p_l} \mathbf{Z}_Q(\tau, \mathbf{x}_l),$$

which has an asymptotic covariance equal to

$$\begin{aligned} &\mathbb{E} \left[\sum_{l=L+1}^M \mathbf{w}_l^{*'} \mathbf{x}_l^* \sqrt{p_l} \mathbf{Z}_Q(\tau, \mathbf{x}_l) \left(\sum_{l'=L+1}^M \mathbf{w}_{l'}^{*'} \mathbf{x}_{l'}^* \sqrt{p_{l'}} \mathbf{Z}_Q(\tau', \mathbf{x}_{l'}) \right)' \right] \\ &= \sum_{l=L+1}^M \sum_{l'=L+1}^M \mathbf{w}_l^{*'} \mathbf{x}_l^* (\min(\tau, \tau') - \tau\tau') \Lambda(\tau, \tau'; \mathbf{x}_l) \cdot \mathbf{1}(l = l') \mathbf{x}_l^{*'} \mathbf{w}_l^* p_l p_{l'} \\ &= (\min(\tau, \tau') - \tau\tau') \mathbb{E}[\mathbf{W}^{*'} \mathbf{X}^* \Lambda(\tau, \tau'; \mathbf{X}) \mathbf{X}^{*'} \mathbf{W}^* | D = 0] \pi_0. \end{aligned}$$

To derive the asymptotic distribution of $\sqrt{N} \left(\hat{\beta}(\cdot; \cdot) - \beta(\cdot; \cdot) \right)$ we note that

$$\begin{aligned} \sqrt{N} \left(\hat{\beta}(\tau; \mathbf{x}_l) - \beta(\tau; \mathbf{x}_l) \right) &= \mathbf{x}_l^{-1} \sqrt{N} \left(\hat{Q}_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) - Q_{\mathbf{Y}|\mathbf{X}}(\tau|\mathbf{x}_l) \right) - \mathbf{x}_l^{-1} \mathbf{w}_l \sqrt{N} \left(\hat{\delta}(\tau) - \delta(\tau) \right) \\ &\xrightarrow{D} \mathbf{x}_l^{-1} \mathbf{Z}_Q(\tau, \mathbf{x}_l) - \mathbf{x}_l^{-1} \mathbf{w}_l \mathbf{Z}_\delta(\tau). \end{aligned} \quad (104)$$

$\mathbf{Z}_Q(\tau, \mathbf{x}_l)$ and $\mathbf{Z}_\delta(\tau)$ are independent processes, since they are computed using disjoint subpopulations (i.e., \mathbf{x}_l for $l = 1, \dots, L$ are not used in the computation of $\hat{\delta}(\tau)$). Therefore, the asymptotic variance of (104) is the sum of the variance of its component terms.

Proof of Theorem 9

Decomposing $\hat{\beta}^M(\tau)$ we see that

$$\sqrt{N} \left(\hat{\beta}^M(\tau) - \bar{\beta}^M(\tau) \right) = \sum_{l=1}^L \beta(\tau; \mathbf{x}_l) \sqrt{N} (\hat{q}_l^M - q_l^M) \quad (105)$$

$$+ \sum_{l=1}^L \sqrt{N} \left(\hat{\beta}(\tau; \mathbf{x}_l) - \beta(\tau; \mathbf{x}_l) \right) \hat{q}_l^M. \quad (106)$$

By a result similar to that in (76) in the main text, term (105) converges to a mean zero Gaussian process with covariance equal to $\frac{\mathbb{C}(\beta(\tau, \mathbf{X}), \beta(\tau', \mathbf{X}) | \mathbf{X} \in \mathbb{X}^M)}{\Pr(\mathbf{X} \in \mathbb{X}^M)}$. Term ((106)) converges to

$$\sum_{l=1}^L \sqrt{N} \left(\hat{\beta}(\tau; \mathbf{x}_l) - \beta(\tau; \mathbf{x}_l) \right) \hat{q}_l^M \xrightarrow{D} \sum_{l=1}^L \mathbf{Z}(\tau, \mathbf{x}_l) q_l^M \quad (107)$$

which has a covariance kernel equal to

$$\begin{aligned} & \mathbb{E} \left[\sum_{l=1}^L \sum_{l'=1}^L \mathbf{Z}(\tau, \mathbf{x}_l) \mathbf{Z}(\tau, \mathbf{x}_{l'}) q_l^M q_{l'}^M \right] \\ &= \mathbb{E} [\mathbf{Z}(\tau, \mathbf{x}_l) \mathbf{Z}(\tau, \mathbf{x}_{l'})] q_l^M q_{l'}^M \\ &= (\min(\tau, \tau') - \tau\tau') \sum_{l=1}^L \sum_{l'=1}^L \frac{\mathbf{x}_l^{-1} \Lambda(\tau, \tau'; \mathbf{x}_l) \mathbf{x}_l^{-1'}}{p^l} \cdot \mathbf{1}(l = l') q_l^M q_{l'}^M \\ &+ \sum_{l=1}^L \sum_{l'=1}^L \mathbf{x}_l^{-1} \mathbf{w}_l \Sigma_\delta(\tau, \tau') \mathbf{w}_{l'}' \mathbf{x}_{l'}^{-1'} q_l^M q_{l'}^M \\ &= \frac{\min(\tau, \tau') - \tau\tau'}{\Pr(\mathbf{X} \in \mathbb{X}^M)} \mathbb{E} [\mathbf{X}^{-1} \Lambda(\tau, \tau', \mathbf{X}) \mathbf{X}^{-1'} | \mathbf{X} \in \mathbb{X}^M] + \sum_{l=1}^L \mathbf{x}_l^{-1} \mathbf{w}_l q_l^M \Sigma_\delta(\tau, \tau') \sum_{l'=1}^L \mathbf{w}_{l'}' \mathbf{x}_{l'}^{-1'} q_{l'}^M \\ &= \Upsilon_1(\tau, \tau') + \Xi_0 \Sigma_\delta(\tau, \tau') \Xi_0'. \end{aligned}$$

Since terms (105) and (106) are uncorrelated, the asymptotic covariance of $\sqrt{N} \left(\hat{\beta}^M(\tau) - \bar{\beta}^M(\tau) \right)$ is equal to the sum of the covariances these two terms.

Proof of Theorem 10

We start by deriving the asymptotic distribution of the sample cumulative distribution function of $\hat{\beta}_p(U; \mathbf{X})$ with U distributed uniformly on $[0, 1]$ independently from \mathbf{X} , while conditioning on $\mathbf{X} \in \mathbb{X}^M$. The CDF

estimand at $c \in \mathbb{R}$ is denoted as $F_{B_p|\mathbf{X} \in \mathbb{X}^M}(c)$ with estimator

$$\begin{aligned}\widehat{F}_{\widehat{\beta}_p(U;\mathbf{X})|\mathbf{X} \in \mathbb{X}^M}(c) &= \frac{\frac{1}{N} \sum_{i=1}^N \left[\int_0^1 \mathbf{1}(\widehat{\beta}_p(u, \mathbf{X}_i) \leq c) du \right] \mathbf{1}(\mathbf{X}_i \in \mathbb{X}^M)}{\frac{1}{N} \sum_{i=1}^N \mathbf{1}(\mathbf{X}_i \in \mathbb{X}^M)} \\ &= \sum_{l=1}^L \left(\int_0^1 \mathbf{1}(\widehat{\beta}_p(u, \mathbf{x}_l) \leq c) du \right) \widehat{q}_l^M.\end{aligned}\quad (108)$$

The integration over $u \in (0, 1)$ can be done exactly since $\widehat{\beta}_p(u, \mathbf{x}_l)$ is piecewise linear for each $l \in \{1, \dots, L\}$ with finitely many pieces. We can express $\widehat{F}_{\widehat{\beta}_p(U;\mathbf{X})|\mathbf{X} \in \mathbb{X}^M}(c)$ as the sum of two terms:

$$\widehat{F}_{\widehat{\beta}_p(U;\mathbf{X})|\mathbf{X} \in \mathbb{X}^M}(c) - F_{B_p|\mathbf{X} \in \mathbb{X}^M}(c) = \sum_{l=1}^L \left(\left[\int_0^1 \mathbf{1}(\widehat{\beta}_p(u, \mathbf{x}_l) \leq c) du \right] - \left[\int_0^1 \mathbf{1}(\beta_p(u, \mathbf{x}_l) \leq c) du \right] \right) \widehat{q}_l^M \quad (109)$$

$$+ \sum_{l=1}^L \left[\int_0^1 \mathbf{1}(\beta_p(u, \mathbf{x}_l) \leq c) du \right] (\widehat{q}_l^M - q_l^M). \quad (110)$$

We will show that these two terms both converge in uniformly over $c \in \mathbb{R}$. For term (109), we have that $\sqrt{N} \left(\widehat{\beta}_p(\tau; \mathbf{x}_l) - \beta_p(\tau; \mathbf{x}_l) \right) \xrightarrow{D} (\mathbf{Z}(\tau, \mathbf{x}_l))_p = \mathbf{Z}_p(\tau, \mathbf{x}_l)$ over $\tau \in (0, 1)$ and all $l = 1, \dots, L$, and $(\cdot)_p$ denotes the p^{th} element of the vector. By the same argument use to show (79) in Appendix B, we have

$$\begin{aligned}& \sqrt{N} \left(\int_0^1 \mathbf{1}(\widehat{\beta}_p(u; \mathbf{x}_l) \leq c) du - \int_0^1 \mathbf{1}(\beta_p(u; \mathbf{x}_l) \leq c) du \right) \\ &= \sqrt{N} \left(\widehat{\beta}_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l); \mathbf{x}_l) - \beta_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l); \mathbf{x}_l) \right) f_{B_p|\mathbf{X}}(c|\mathbf{x}_l) + o_p(1) \\ &\xrightarrow{D} \mathbf{Z}_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l), \mathbf{x}_l) f_{B_p|\mathbf{X}}(c|\mathbf{x}_l).\end{aligned}$$

This convergence is uniform in $c \in \mathbb{R}$ since $F_{B_p|\mathbf{X}}(c|\mathbf{x}_l)$ ranges between 0 and 1, and uniform in \mathbf{x}_l since its support is finite. Therefore,

$$\sum_{l=1}^L \sqrt{N} \left(\int_0^1 \mathbf{1}(\widehat{\beta}_p(u; \mathbf{x}_l) \leq c) du - \int_0^1 \mathbf{1}(\beta_p(u; \mathbf{x}_l) \leq c) du \right) \widehat{q}_l^M \xrightarrow{D} \sum_{l=1}^L \mathbf{Z}_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l), \mathbf{x}_l) f_{B_p|\mathbf{X}}(c|\mathbf{x}_l) q_l^M \quad (111)$$

for $c \in \mathbb{R}$. Also, (110) will converge over $c \in \mathbb{R}$ to a Gaussian process $\mathbf{Z}_{2p}(c)$ with asymptotic covariance of

$$\mathbb{E} [\mathbf{Z}_{2p}(c) \mathbf{Z}_{2p}(c)'] = \frac{\mathbb{C} (F_{B_p|\mathbf{X}}(c|\mathbf{X}), F_{B_p|\mathbf{X}}(c'|\mathbf{X}) | \mathbf{X} \in \mathbb{X}^M)}{\Pr(\mathbf{X} \in \mathbb{X}^M)}.$$

Note that $\mathbf{Z}_{2p}(c)$ and $\sum_{l=1}^L \mathbf{Z}_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l), \mathbf{x}_l) f_{B_p|\mathbf{X}}(c|\mathbf{x}_l) q_l^M$ are uncorrelated since the variation in the latter is conditional on \mathbf{X} while that in the former depends on \mathbf{X} only. Therefore,

$$\sqrt{N} \left(\widehat{F}_{\widehat{\beta}_p(U;\mathbf{X})|\mathbf{X} \in \mathbb{X}^M}(c) - F_{B_p|\mathbf{X} \in \mathbb{X}^M}(c) \right) \xrightarrow{D} \sum_{l=1}^L \mathbf{Z}_p(F_{B_p|\mathbf{X}}(c|\mathbf{x}_l), \mathbf{x}_l) f_{B_p|\mathbf{X}}(c|\mathbf{x}_l) q_l^M + \mathbf{Z}_{2p}(c)$$

for $c \in \mathbb{R}$.

Using the same invertibility argument as in (82), we see that

$$\begin{aligned} \sqrt{N} \left(\hat{\beta}_p^M(\tau) - \beta_p^M(\tau) \right) &\xrightarrow{D} \frac{\sum_{l=1}^L \mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), \mathbf{x}_l) f_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l) q_l^M + \mathbf{Z}_{2p}(\beta_p^M(\tau))}{f_{B_p|\mathbf{X} \in \mathbb{X}^M}(\beta_p^M(\tau))} \\ &= \mathbf{Z}_{\beta_p}(\tau) \end{aligned} \quad (112)$$

uniformly over $\tau \in (0, 1)$.

To conclude this proof, we evaluate $\mathbb{E} [\mathbf{Z}_{\beta_p}(\tau) \mathbf{Z}_{\beta_p}(\tau')']$, the asymptotic covariance of ((112)):

$$\begin{aligned} \mathbb{E} \left[\mathbf{Z}_{\beta_p}(\tau) \mathbf{Z}_{\beta_p}(\tau')' \right] &= \frac{\sum_{l=1}^L \sum_{l'=1}^L f_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l) f_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}) q_l^M q_{l'}^M}{f_{B_p|\mathbf{X} \in \mathbb{X}^M}(\beta_p^M(\tau)) f_{B_p|\mathbf{X} \in \mathbb{X}^M}(\beta_p^M(\tau'))} \\ &\quad \times \mathbb{E} \left[\mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), \mathbf{x}_l) \mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}), \mathbf{x}_{l'}) \right] \\ &\quad + \frac{\mathbb{E} \left[\mathbf{Z}_{2p}(\beta_p^M(\tau)) \mathbf{Z}_{2p}(\beta_p^M(\tau')) \right]}{f_{B_p|\mathbf{X} \in \mathbb{X}^M}(\beta_p^M(\tau)) f_{B_p|\mathbf{X} \in \mathbb{X}^M}(\beta_p^M(\tau'))}, \end{aligned}$$

where

$$\begin{aligned} &\mathbb{E} \left[\mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), \mathbf{x}_l) \mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}), \mathbf{x}_{l'}) \right] \\ &= (\min(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'})) - F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l) F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'})) \\ &\quad \times e_p' \frac{\mathbf{x}_l^{-1} \Lambda(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}); \mathbf{x}_l) \mathbf{x}_l^{-1l'}}{p_l} e_p \cdot \mathbf{1}(l = l') \\ &\quad + e_p' \mathbf{x}_l^{-1} \mathbf{w}_l \Sigma_\delta(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'})) \mathbf{w}_l' \mathbf{x}_l^{-1l'} e_p, \end{aligned}$$

and

$$\begin{aligned} &\sum_{l=1}^L \sum_{l'=1}^L f_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l) q_l^M \mathbb{E} \left[\mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{x}_l), \mathbf{x}_l) \mathbf{Z}_p(F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}), \mathbf{x}_{l'}) \right] f_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{x}_{l'}) q_{l'}^M \\ &= \mathbb{E} \left[\frac{(\min(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{X})) - F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}) F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{X}))}{\Pr(\mathbf{X} \in \mathbb{X}^M)} \right] \\ &\quad \times e_p' \mathbf{X}^{-1} \Lambda(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{X}); \mathbf{X}) \mathbf{X}^{-1l'} e_p f_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}) f_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{X}) | \mathbf{X} \in \mathbb{X}^M \\ &\quad + e_p' \mathbb{E} \left[f_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}) f_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\tilde{\mathbf{X}}) \mathbf{X}^{-1} \mathbf{W} \Sigma_\delta(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\tilde{\mathbf{X}})) \right. \\ &\quad \left. \times \tilde{\mathbf{W}}' \tilde{\mathbf{X}}^{-1l'} | \mathbf{X} \in \mathbb{X}^M, \tilde{\mathbf{X}} \in \mathbb{X}^M \right] e_p \\ &= \Upsilon_3(\tau, \tau') + \Upsilon_4(\tau, \tau'), \end{aligned} \quad (113)$$

with $\tilde{\mathbf{X}}$ is an independent copy of \mathbf{X} . Finally,

$$\begin{aligned} \mathbb{E} \left[\mathbf{Z}_{2p}(\beta_p^M(\tau)) \mathbf{Z}_{2p}(\beta_p^M(\tau')) \right] &= \frac{\mathbb{C}(F_{B_p|\mathbf{X}}(\beta_p^M(\tau)|\mathbf{X}), F_{B_p|\mathbf{X}}(\beta_p^M(\tau')|\mathbf{X}) | \mathbf{X} \in \mathbb{X}^M)}{\Pr(\mathbf{X} \in \mathbb{X}^M)} \\ &= \Upsilon_2(\tau, \tau'). \end{aligned}$$