Supplement to "Semiparametric Quantile Regression Imputation for a Complex Survey with Application to the Conservation Effects Assessment Project"

Emily Berg and Cindy Yu

A: Regularity Conditions

Regularity conditions required for Lemma 2 and Theorem 1 are as follows.

1. $\boldsymbol{\theta}_o$ is the unique value such that $E[\boldsymbol{g}_i(y_i, \boldsymbol{\theta}_o)] = \mathbf{0}$.

2. The function $\boldsymbol{g}_i(y_i, \boldsymbol{\theta})$ is twice differentiable with respect to both y_i, and $\boldsymbol{\theta}$, and the second partial derivatives are continuous.

3. The function $\boldsymbol{g}_i(y_i, \boldsymbol{\theta})$ has bounded $2+\delta$ moments, and the first two derivatives of $\boldsymbol{g}_i$ have bounded $2+\delta$ moments. More precisely,

$$E_{\boldsymbol{\theta}_o}[|h_i(y_i, \boldsymbol{\theta})|^{2+\delta}] < M, \tag{45}$$

where $h_i(y_i, \boldsymbol{\theta})$ denotes an element of $\boldsymbol{g}_i$ or a first or second partial derivative of $\boldsymbol{g}_i(y_i; \boldsymbol{\theta})$ with respect to y_i or $\boldsymbol{\theta}$.

4. Assume that $\bar{h}_{HT}(\boldsymbol{\theta}) - \bar{h}_N(\boldsymbol{\theta})$ converges to 0 uniformly in $\boldsymbol{\theta}$, where

$$\begin{aligned} \bar{h}_{HT}(\boldsymbol{\theta}) &:= \bar{h}_{HT,N}(\boldsymbol{\theta}) = N^{-1}\sum_{i=1}^{N} \pi_i^{-1} I_i h_i(y_i, \boldsymbol{\theta}), \qquad (46) \\ \bar{h}_N(\boldsymbol{\theta}) &= N^{-1}\sum_{i=1}^{N} h_i(y_i, \boldsymbol{\theta}), \end{aligned}$$

and $h_i(y_i, \boldsymbol{\theta})$ has the same interpretation as in condition 7 above. This condition means that for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$P(|\bar{h}_{HT}(\boldsymbol{\theta}) - \bar{h}_N(\boldsymbol{\theta})| > \epsilon) < \delta, \tag{47}$$

for all N greater than some value M, and for all $\boldsymbol{\theta}$. See Kim and Park (2010) for a similar condition.

5. $||\boldsymbol{G}_n(\hat{\boldsymbol{\theta}})|| \leq o_p(n^{-0.5}) + \inf_{\boldsymbol{\theta}} ||\boldsymbol{G}_n(\boldsymbol{\theta})||$.

B. Proofs

B.1 Proof of Lemma 1

Consider the objective function,

$$\begin{aligned} Z_{HT}(\Delta) &= \frac{1}{K_n}\left\{\sum_{i=1}^{n} nN^{-1}\pi_i^{-1}\delta_i\left[\rho_\tau\left(u_{i\tau} - \sqrt{\frac{K_n}{n}}\boldsymbol{B}(x_i)'\Delta\right) - \rho_\tau(u_{i\tau})\right]\right\} \\ &+\frac{1}{K_n}\left\{\frac{\tilde{\lambda}_n}{2}\left(\boldsymbol{\beta}_\tau^* + \sqrt{\frac{K_n}{n}}\Delta\right)'\boldsymbol{D}_m'\boldsymbol{D}_m\left(\boldsymbol{\beta}_\tau^* + \sqrt{\frac{K_n}{n}}\Delta\right) - \frac{K_n\tilde{\lambda}_n}{2}(\boldsymbol{\beta}_\tau^*)'\boldsymbol{D}_m'\boldsymbol{D}_m(\boldsymbol{\beta}_\tau^*)\right\}, \end{aligned} \tag{48}$$

where $\tilde{\lambda}_n = \lambda_n n\hat{N}N^{-1}$, $u_{i\tau} = y_i - \boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^*$ and $\boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^*$ is the L_∞ approximation to $q_\tau(x)$, the true quantile function in the superpopulation. The objective function is multiplied by nN^{-1} because for a simple random sample $nN^{-1}\pi_i^{-1} = 1$, and the resulting objective function simplifies to a simple sum. By the definition of $\hat{\boldsymbol{\beta}}_\tau$, the minimizer of $Z_{HT}(\Delta)$ is,

$$\sqrt{\frac{n}{K_n}}\left(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau^*\right). \tag{49}$$

From Knight's identity (Knight, 1998),

$$\begin{aligned} Z_{HT}(\Delta) &= \left\{\frac{1}{K_n}Z_{1,HT} + \tilde{\lambda}_n\sqrt{\frac{1}{K_n n}}(\boldsymbol{\beta}_\tau^*)'\boldsymbol{D}_m'\boldsymbol{D}_m\right\}\Delta \\ & \frac{1}{K_n}Z_{2,HT}(\Delta) + \frac{\tilde{\lambda}_n}{2n}\Delta'\boldsymbol{D}_m'\boldsymbol{D}_m\Delta, \end{aligned} \tag{50}$$

where

$$\begin{aligned} Z_{1,HT} &= -\sqrt{\frac{K_n}{n}}\sum_{i=1}^{n} nN^{-1}\pi_i^{-1}\delta_i\boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau}), \\ Z_{2,HT}(\Delta) &= \sum_{i=1}^{n} nN^{-1}\delta_i\pi_i^{-1}\int_0^{\nu_{ni}}(I[u_i \leq s] - I[u_i \leq 0])ds, \end{aligned} \tag{51}$$

and $\nu_{ni} = K_n^{0.5} n^{-0.5} \boldsymbol{B}(x_i)'\Delta$.

We first consider $Z_{2,HT}(\Delta)$ and follow the approach of Koenker (2004) and Yoshida (2013). Let $q_{\tau i}$ satisfy $F_{y|x,i}(q_{\tau i} \,|\, x_i) = P(y_i \leq q_{\tau i} \,|\, x_i) = \tau$, and let $f_{y|x,i}(t)$ denote the derivative of $F_{y|x,i}(\cdot)$ evaluated at t. Let $\boldsymbol{x}_N$ be the vector of x_i for $i = 1, \ldots, N$. Then,

$$
\begin{aligned}
E[Z_{2,HT}(\Delta) \,|\, \boldsymbol{x}_N] &= E\left[\frac{n}{N}\sum_{i=1}^{N} \delta_i \int_0^{\nu_{ni}} I[0 \leq u_i \leq s] ds \,|\, \boldsymbol{x}_N\right] \qquad (52)\\
&= \frac{K_n}{N}\sum_{i=1}^{N} \delta_i \int_0^{\boldsymbol{B}(x_i)'\Delta} \sqrt{\frac{n}{K_n}} \left[F_{y|x,i}\left(\boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^* + t\sqrt{\frac{K_n}{n}}\right) - F_{y|x,i}\left(\boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^*\right)\right] dt\\
&= \frac{K_n}{N}\sum_{i=1}^{N} \int_0^{\boldsymbol{B}(x_i)'\Delta} f_{y|x,i}(\boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^*) t dt + o(1)\\
&= \frac{K_n}{2N}\sum_{i=1}^{N} \delta_k \Delta' \boldsymbol{B}(x_i) f_{y|x,i}(\boldsymbol{B}(x_i)'\boldsymbol{\beta}_\tau^*) \boldsymbol{B}(x_i)'\Delta\\
&= \frac{K_n}{2N}\sum_{i=1}^{N} \delta_k \Delta' \boldsymbol{B}(x_i) f_{y|x,i}(q_{\tau i} + o(1)) \boldsymbol{B}(x_i)'\Delta\\
&= \frac{K_n}{2}\Delta' \boldsymbol{H}(\tau)\Delta + o_p(1),
\end{aligned}
$$

where $\boldsymbol{H}(\tau) = E[p_i \boldsymbol{B}(x_i) f_{y|x,i}(q_{\tau i}) \boldsymbol{B}(x_i)']$. This shows that $E[Z_{2,HT}(\Delta) \,|\, \boldsymbol{x}_N] = O_p(K_n)$. We next show that $V\{Z_{2,HT}(\Delta)\,|\,\boldsymbol{x}_N\} = o_p(K_n)$. Using the argument of double expectations,

$$\begin{aligned}
V\{Z_{2,HT}(\Delta)\,|\,\boldsymbol{x}_N\} &= E[V\{Z_{2,HT}(\Delta)\,|\,\mathscr{F}_N,\boldsymbol{x}_N\}\,|\,\boldsymbol{x}_N] + V\{E[Z_{2,HT}(\Delta)\,|\,\mathscr{F}_N,\boldsymbol{x}_N]\,|\,\boldsymbol{x}_N\} \\
&= E\left\{\frac{n^2}{N^2}\sum_{i=1}^N \delta_i(\pi_i^{-1}-1)(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)^2\right\} \\
&\quad + E\left\{\frac{n^2}{N^2}\sum_{i=1}^N\sum_{j\neq i}\delta_i\delta_j\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)(\int_0^{\nu_{n_j}} I[0\le u_j\le s]ds)\right\} \\
&\quad + V\left\{nN^{-1}\sum_{i=1}^N\delta_i(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)\right\} \\
&\le E\left\{\frac{n}{N}K_2\sum_{i=1}^N\delta_i(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)\sqrt{\frac{K_n}{n}}\max\{||\boldsymbol{B}(x_i)||||\Delta||\}\right\} \\
&\quad + E\left\{\frac{n^2}{N^2}\sum_{i=1}^N\sum_{j\neq i}\delta_i\delta_j n_B^{-1}K_3(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)(\int_0^{\nu_{n_j}} I[0\le u_j\le s]ds)\right\} \\
&\quad + n^2N^{-2}\sum_{i=1}^N\delta_i E[(\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds)^2] \\
&= o(K_n).
\end{aligned} \tag{53}$$

The last equality holds because $E[nN^{-1}\sum_{i=1}^N\delta_i\int_0^{\nu_{n_i}} I[0\le u_i\le s]ds] = O(K_n)$, $||\boldsymbol{B}(x_i)||$ is bounded and $K_n = o(n^{0.5})$.

Then,

$$Z_{HT}(\Delta) = Q_{HT}(\Delta) + o_p(1), \tag{54}$$

where

$$\begin{aligned}
Q_{HT}(\Delta) &= \left\{-\sqrt{\frac{1}{K_n n}}\sum_{i=1}^n\frac{n}{N}\pi_i^{-1}\delta_i\boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau}) + \tilde{\lambda}_n\sqrt{\frac{1}{K_n n}}(\boldsymbol{\beta}_\tau^*)'\boldsymbol{D}_m'\boldsymbol{D}_m\right\}\Delta \\
&\quad + \Delta'\left[\frac{1}{2}\boldsymbol{H}(\tau) + \frac{\tilde{\lambda}_n{}'}{2n}\boldsymbol{D}_m'\boldsymbol{D}_m\right]\Delta.
\end{aligned} \tag{55}$$

The $Z_{HT}(\Delta)$ is quadratic in Δ and is minimized at $\hat{\Delta}$, where

$$\hat{\Delta} = \sqrt{\frac{1}{K_n n}}\left[\boldsymbol{H}(\tau) + \frac{\tilde{\lambda}_n}{n}\boldsymbol{D}_m'\boldsymbol{D}_m\right]^{-1}\left\{\sum_{i=1}^n nN^{-1}\pi_i^{-1}\delta_i\boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau}) - \tilde{\lambda}_n\boldsymbol{D}_m'\boldsymbol{D}_m(\boldsymbol{\beta}_\tau^*)\right\}. \tag{56}$$

The asymptotic distribution of $\sqrt{\frac{n}{K_n}}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}^*_\tau)$ is the same as the asymptotic distribution of $\hat{\Delta}$.

We next consider the asymptotic distribution of $\hat{\Delta}$. By the proof of lemma 1 in Yoshida (2013),

$$E[|nN^{-1}\pi_i^{-1}\delta_i \boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau})|^{2+\delta}] < \infty, \tag{57}$$

and by (21),

$$\sqrt{n}(\bar{w}_{HT}(\tau) - \bar{w}_N(\tau)) \,|\, \mathscr{F}_N \xrightarrow{d} N(0, \boldsymbol{V}_{1,\infty}(\tau)) \quad a.s., \tag{58}$$

where

$$\begin{aligned} \bar{w}_{HT}(\tau) &= N^{-1}\sum_{i=1}^{n} \pi_i^{-1}\delta_i \boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau}), \\ \bar{w}_N(\tau) &= N^{-1}\sum_{i=1}^{N} \delta_i \boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau}), \end{aligned} \tag{59}$$

and

$$\boldsymbol{V}_{1,\infty}(\tau) = \lim_{N\to\infty} \frac{n}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j}\delta_i\delta_j \boldsymbol{B}(x_i)'\psi_\tau(u_{i\tau})\boldsymbol{B}(x_j)'\psi_\tau(u_{j\tau}). \tag{60}$$

By the proof of Chen and Yu (2016),

$$\sqrt{n}\bar{w}_N(\tau) \xrightarrow{d} N(0, f_\infty \tau(1-\tau)\boldsymbol{\Phi}), \tag{61}$$

where $\Phi = E[p_i \boldsymbol{B}(x_i)\boldsymbol{B}(x_i)']$. By Theorem 1.3.6 from Fuller (2011),

$$\sqrt{n}\bar{w}_{HT}(\tau) \xrightarrow{d} N\left(0, \boldsymbol{V}_{1,\infty}(\tau) + f_\infty\tau(1-\tau)\Phi\right). \tag{62}$$

From the above expansion,

$$\sqrt{\frac{n}{K_n}}\left(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}^*_\tau\right) \stackrel{d}{=} \hat{\Delta} = \sqrt{\frac{1}{K_n}}\boldsymbol{\Omega}_n(\tau)^{-1}\sqrt{n}\bar{w}_{HT}(\tau) - \frac{\tilde{\lambda}_n}{\sqrt{K_n n}}\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{D}'_m\boldsymbol{D}_m\boldsymbol{\beta}^*_\tau, \tag{63}$$

where $\boldsymbol{\Omega}_n(\tau) = [\boldsymbol{H}(\tau) + \tilde{\lambda}_n n^{-1} \boldsymbol{D}'_m \boldsymbol{D}_m]$. Then,

$$\sqrt{\frac{n}{K_n}} \left(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}^*_\tau + \frac{\tilde{\lambda}_n}{n} \boldsymbol{\Omega}_n(\tau)^{-1} \boldsymbol{D}'_m \boldsymbol{D}_m \boldsymbol{\beta}^*_\tau \right) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_\infty(\tau)), \tag{64}$$

where

$$\boldsymbol{\Sigma}_\infty(\tau) = lim_{N \to \infty} \frac{1}{K_n} \boldsymbol{\Omega}_n(\tau)^{-1} \left(\boldsymbol{V}_{1,\infty}(\tau) + f_\infty \tau (1-\tau) \Phi \right) \boldsymbol{\Omega}_n(\tau)^{-1}. \tag{65}$$

Result (25) of lemma 1 follows because $\hat{q}_\tau(x_i) = \boldsymbol{B}(x_i)' \hat{\boldsymbol{\beta}}_\tau$, and (26) follows from the result of Barrow and Smith (1978) provided in (23).

B.2 Proof of Lemma 2

The estimator (17), evaluated at $\boldsymbol{\theta}_o$ is

$$\sqrt{n} \boldsymbol{G}_n(\boldsymbol{\theta}_o) \quad = \quad \frac{\sqrt{n}}{N} \sum_{i=1}^{n} \pi_i^{-1} \delta_i \boldsymbol{g}_i(y_i, \boldsymbol{\theta}_o) + \frac{\sqrt{n}}{N} \sum_{i=1}^{n} \pi_i^{-1} (1-\delta_i) \int_0^1 \boldsymbol{g}_i(\boldsymbol{B}(x_i)' \hat{\boldsymbol{\beta}}_\tau, \boldsymbol{\theta}_o) d\tau \tag{66}$$

To simplify notation, we write,

$$\sqrt{n} \boldsymbol{G}_n(\boldsymbol{\theta}_o) = \sqrt{n} \boldsymbol{G}_{obs}(\boldsymbol{\theta}_o) + \sqrt{n} \boldsymbol{G}_{mis}(\boldsymbol{\theta}_o), \tag{67}$$

where

$$\sqrt{n} \boldsymbol{G}_{obs}(\boldsymbol{\theta}_o) \quad = \quad \frac{\sqrt{n}}{N} \sum_{i=1}^{n} \pi_i^{-1} \delta_i \boldsymbol{g}_i(y_i, \boldsymbol{\theta}_o), \tag{68}$$

and

$$\sqrt{n} \boldsymbol{G}_{mis}(\boldsymbol{\theta}_o) \quad = \quad \frac{\sqrt{n}}{N} \sum_{i=1}^{n} \pi_i^{-1} (1-\delta_i) \int_0^1 \boldsymbol{g}_i(\boldsymbol{B}(x_i)' \hat{\boldsymbol{\beta}}_\tau, \boldsymbol{\theta}_o) d\tau. \tag{69}$$

Consider $\boldsymbol{G}_{mis}(\boldsymbol{\theta}_o)$. By the mean value theorem,

$$\begin{aligned}\boldsymbol{G}_{mis}(\boldsymbol{\theta}_o) &= \frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 \boldsymbol{g}(q_\tau(x_i);\boldsymbol{\theta}_o)d\tau \\ &+\frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 [\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)(\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau - q_\tau(x_i))d\tau] \\ &+\frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 [\ddot{\boldsymbol{g}}_{i,y}(\tilde{z}_i;\boldsymbol{\theta}_o)(\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau - q_\tau(x_i))^2 d\tau],\end{aligned} \tag{70}$$

where $\tilde{z}_i$ is between $\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau$ and $q_\tau(x_i)$, and $\dot{\boldsymbol{g}}_{i,y}$ and $\ddot{\boldsymbol{g}}_{i,y}$ are the vectors of first and second derivatives of the elements of $\boldsymbol{g}_i$ with respect to y_i.

By the proof of lemma 1,

$$\begin{aligned}\sqrt{n}(\hat{q}_\tau(x_i) - q_\tau(x_i)) &= \frac{1}{\sqrt{n}}\boldsymbol{B}(x_i)'\boldsymbol{\Omega}_n(\tau)^{-1}\sum_{j=1}^{n} nN^{-1}\pi_j^{-1}\delta_j\boldsymbol{B}(x_j)\psi_\tau(u_{j\tau}) \\ &-\frac{\tilde{\lambda}}{\sqrt{n}}\boldsymbol{B}(x_i)'\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{D}_m'\boldsymbol{D}_m\boldsymbol{\beta}_\tau^* - \sqrt{n}b_\tau^a(x_i) + o_p(1).\end{aligned} \tag{71}$$

Therefore,

$$\sqrt{n}\boldsymbol{G}_{mis}(\boldsymbol{\theta}_o) = \boldsymbol{G}_{1,mis}(\boldsymbol{\theta}_o) + \boldsymbol{G}_{2,mis}(\boldsymbol{\theta}_o) + \boldsymbol{G}_{3,mis}(\boldsymbol{\theta}_o) + \boldsymbol{G}_{4,mis}(\boldsymbol{\theta}_o) + \boldsymbol{G}_{5,mis}(\boldsymbol{\theta}_o), \tag{72}$$

where

$$\begin{aligned}\boldsymbol{G}_{1,mis}(\boldsymbol{\theta}_o) &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 \boldsymbol{g}_i(q_\tau(x_i);\boldsymbol{\theta}_o)d\tau, \\ \boldsymbol{G}_{2,mis}(\boldsymbol{\theta}_o) &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 \dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)'\boldsymbol{\Omega}_n(\tau)^{-1}\frac{1}{N}\sum_{j=1}^{n}\pi_j^{-1}\delta_j\boldsymbol{B}(x_j)\psi_\tau(u_{j\tau})d\tau \\ \boldsymbol{G}_{3,mis}(\boldsymbol{\theta}_o) &= -\frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 \dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\frac{\tilde{\lambda}}{n}\boldsymbol{B}(x_i)'\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{D}_m'\boldsymbol{D}_m\boldsymbol{\beta}_\tau^* \\ \boldsymbol{G}_{4,mis}(\boldsymbol{\theta}_o) &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1 \dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)b_\tau^a(x_i)d\tau,\end{aligned} \tag{73}$$

and

$$\boldsymbol{G}_{5,mis}(\boldsymbol{\theta}_o) = \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\int_0^1 \ddot{\boldsymbol{g}}_{i,y}(\tilde{z}_i;\boldsymbol{\theta}_o)(\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau - q_\tau(x_i))^2 d\tau. \tag{74}$$

As in Chen and Yu (2016),

$$\boldsymbol{G}_{3,mis}(\boldsymbol{\theta}_o) = O_p(n^{0.5}K_n^{-(p+2)}) = O_p(n^{-\frac{1}{2(2p+3)}}) = o_p(1), \tag{75}$$

and

$$\boldsymbol{G}_{4,mis}(\boldsymbol{\theta}_o) = O_p(n^{0.5}K_n^{-(p+2)}) = O_p(n^{-\frac{1}{2(2p+3)}}) = o_p(1). \tag{76}$$

Exchanging the order of integration and summation, $\boldsymbol{G}_{2,mis}(\boldsymbol{\theta}_o)$ can be expressed as,

$$\boldsymbol{G}_{2,mis}(\boldsymbol{\theta}_o) = \int_0^1 \frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)'\boldsymbol{\Omega}_n(\tau)^{-1}\frac{\sqrt{n}}{N}\sum_{j=1}^{n}\pi_j^{-1}\delta_j\boldsymbol{B}(x_j)\psi_\tau(u_{j\tau})d\tau.$$

By the assumptions (21) and $E[||\dot{\boldsymbol{g}}_{i,y}(Y;\boldsymbol{\theta}_o)||^{2+\delta}] < \infty$,

$$\begin{aligned}\boldsymbol{G}_{2,mis}(\boldsymbol{\theta}_o) &= \frac{1}{\sqrt{n}}\int_0^1 E[(1-p_i)\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)']\boldsymbol{\Omega}_n(\tau)^{-1}\sum_{j=1}^{n}\pi_j^{-1}\frac{n}{N}\delta_j\boldsymbol{B}(x_j)\psi_\tau(u_{j\tau})d\tau + o_p(1)\\ &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}\delta_i\int_0^1 E[(1-p_i)\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)']\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{B}(x_i)\psi_\tau(u_{i\tau})d\tau + o_p(1).\end{aligned}$$

Similarly,

$$\begin{aligned}\boldsymbol{G}_{5,mis}(\boldsymbol{\theta}_o) &= \int_0^1 \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)[\ddot{\boldsymbol{g}}_{i,y}(\tilde{z}_i;\boldsymbol{\theta}_o)(\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau - q_\tau(x_i))^2]d\tau\\ &= O_p(\frac{1}{\sqrt{n}}).\end{aligned}$$

Combining above terms,

$$\begin{aligned}\sqrt{n}\boldsymbol{G}(\boldsymbol{\theta}_o) &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}\delta_i\boldsymbol{g}_i(y_i;\boldsymbol{\theta}_o) \\ &\quad +\frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}(1-\delta_i)\int_0^1\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)d\tau \\ &\quad +\frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}\delta_i\int_0^1 E[(1-p_i)\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)']\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{B}(x_i)\psi_\tau(u_{i\tau})d\tau \\ &\quad +o_p(1) \\ &= \frac{\sqrt{n}}{N}\sum_{i=1}^{n}\pi_i^{-1}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)+o_p(1),\end{aligned} \tag{77}$$

where

$$\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)=\delta_i\boldsymbol{g}_i(y_i;\boldsymbol{\theta}_o)+(1-\delta_i)\int_0^1\boldsymbol{g}_i(q_\tau(x_i);\boldsymbol{\theta}_o)d\tau+\delta_i\boldsymbol{h}_{ni}(\boldsymbol{\theta}_o), \tag{78}$$

and

$$\boldsymbol{h}_{n_i}(\boldsymbol{\theta}_o)=\int_0^1 E[(1-p_i)\dot{\boldsymbol{g}}_{i,y}(q_\tau(x_i);\boldsymbol{\theta}_o)\boldsymbol{B}(x_i)']\boldsymbol{\Omega}_n(\tau)^{-1}\boldsymbol{B}(x_i)\psi_\tau(u_{i\tau})d\tau. \tag{79}$$

From assumption (21),

$$\sqrt{n}\left(\frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)-\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)\right)\xrightarrow{d}N(\mathbf{0},\boldsymbol{V}_{\xi,N}(\boldsymbol{\theta}_o))\quad a.s., \tag{80}$$

where

$$\boldsymbol{V}_{\xi,N}=nN^{-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)\boldsymbol{\xi}_j(\boldsymbol{\theta}_o)'. \tag{81}$$

By the proof of Chen and Yu (2016),

$$\frac{\sqrt{n}}{N}\left(\sum_{i=1}^{N}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)\right)\xrightarrow{d}N(0,f_\infty V\{\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)\}). \tag{82}$$

An application of Theorem 1.3.6 of Fuller (2011) gives,

$$\sqrt{n}(\frac{1}{N}\sum_{i=1}^{n}\pi_i^{-1}\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)) \xrightarrow{d} N(0, f_\infty V\{\boldsymbol{\xi}_i(\boldsymbol{\theta}_o)\} + \lim_{N\to\infty}\boldsymbol{V}_{\xi,N}(\boldsymbol{\theta}_o))'. \tag{83}$$

B.3 Proof of Theorem 1

First, we prove consistency of $\hat{\boldsymbol{\theta}}$ using an approach analogous to Chen and Yu (2016). We show that the following two conditions of Pakes and Pollard (1989) hold.

1. $\sup_\theta|\boldsymbol{G}_n(\boldsymbol{\theta}) - \boldsymbol{G}(\boldsymbol{\theta})| = o_p(1)$

2. For every sequence of real numbers $\zeta_n \to 0$,

$$\sup_{|\boldsymbol{\theta}-\boldsymbol{\theta}_o|<\zeta_n}|\boldsymbol{G}_n(\boldsymbol{\theta}) - \boldsymbol{G}(\boldsymbol{\theta}) - \boldsymbol{G}_n(\boldsymbol{\theta}_o)| = o_p(n^{-0.5}), \tag{84}$$

where

$$\boldsymbol{G}(\boldsymbol{\theta}) = E[\boldsymbol{G}_N(\boldsymbol{\theta}, \boldsymbol{y})], \tag{85}$$

and

$$\boldsymbol{G}_N(\boldsymbol{\theta}, \boldsymbol{y}) = N^{-1}\sum_{i=1}^{N}\delta_i g(y_i, x_i, \boldsymbol{\theta}). \tag{86}$$

As in Chen and Yu (2016), $\boldsymbol{G}_n(\boldsymbol{\theta})$ can be decomposed as,

$$\boldsymbol{G}_n(\boldsymbol{\theta}) = B_{1,HT}(\boldsymbol{\theta}) + B_{2,HT}(\boldsymbol{\theta}) + B_{3,HT}(\boldsymbol{\theta}), \tag{87}$$

where

$$\begin{aligned}
B_{1,HT}(\boldsymbol{\theta}) &= N^{-1}\sum_{i=1}^{N} I_i\pi_i^{-1}\boldsymbol{g}_i(y_i,\boldsymbol{\theta}), &(88)\\
B_{2,HT}(\boldsymbol{\theta}) &= N^{-1}\sum_{i=1}^{N} I_i\pi_i^{-1}(1-\delta_i)(\boldsymbol{\mu}_{g|x}(x_i,\boldsymbol{\theta})-\boldsymbol{g}_i(y_i,\boldsymbol{\theta}))\\
B_{3,HT}(\boldsymbol{\theta}) &= N^{-1}\sum_{i=1}^{N} I_i\pi_i^{-1}(1-\delta_i)(\hat{\boldsymbol{\mu}}_{g|x}(x_i,\boldsymbol{\theta})-\boldsymbol{\mu}_{g|x}(x_i,\boldsymbol{\theta})),\\
\hat{\boldsymbol{\mu}}_{g|x}(x_i,\boldsymbol{\theta}) &= \int_0^1 \boldsymbol{g}_i(\boldsymbol{B}(x_i)'\hat{\boldsymbol{\beta}}_\tau,\boldsymbol{\theta})d\tau,
\end{aligned}$$

and

$$\boldsymbol{\mu}_{g|x}(x_i,\boldsymbol{\theta}) = \int_0^1 \boldsymbol{g}_i(q_\tau(x_i),\boldsymbol{\theta})d\tau. \qquad (89)$$

From the proof of Lemma 2,

$$B_{3,HT}(\boldsymbol{\theta}) = N^{-1}\sum_{i=1}^{N}\pi_i^{-1}I_i\delta_i\boldsymbol{h}_{ni}(\boldsymbol{\theta})+o_p(n^{-0.5}). \qquad (90)$$

Using this decomposition,

$$\begin{aligned}
\boldsymbol{G}_n(\boldsymbol{\theta})-\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y}) &= N^{-1}\sum_{i=1}^{N} I_i\pi_i^{-1}\delta_i\boldsymbol{g}_i(y_i,\boldsymbol{\theta})-N^{-1}\sum_{i=1}^{N}\delta_i\boldsymbol{g}_i(y_i,\boldsymbol{\theta}) &(91)\\
&\quad +B_{2,HT}(\boldsymbol{\theta})+N^{-1}\sum_{i=1}^{N}\pi_i^{-1}I_i\delta_i\boldsymbol{h}_{ni}(\boldsymbol{\theta})+o_p(n^{-0.5}).
\end{aligned}$$

By (21),

$$N^{-1}\sum_{i=1}^{N} I_i\pi_i^{-1}\boldsymbol{g}_i(y_i,\boldsymbol{\theta})-N^{-1}\sum_{i=1}^{N}\boldsymbol{g}_i(y_i,\boldsymbol{\theta}) = o_p(1). \qquad (92)$$

Similarly,

$$B_{2,HT}(\boldsymbol{\theta}) = N^{-1}\sum_{i=1}^{N}(1-\delta_i)(\boldsymbol{\mu}_{g|x}(x_i,\boldsymbol{\theta})-\boldsymbol{g}_i(y_i,\boldsymbol{\theta}))+O_p(n^{-0.5}) \quad a.s., \qquad (93)$$

and by the law of large numbers,

$$N^{-1}\sum_{i=1}^{N}(1-\delta_i)(\boldsymbol{\mu}_{g|x}(x_i,\boldsymbol{\theta})-\boldsymbol{g}_i(y_i,\boldsymbol{\theta}))=o_p(1), \tag{94}$$

so that $B_{2,HT}(\boldsymbol{\theta})=o_p(1)$. Because $E[\boldsymbol{h}_{ni}(\boldsymbol{\theta})]=o_p(1)$,

$$\begin{aligned} B_{3,HT}(\boldsymbol{\theta}) &= N^{-1}\sum_{i=1}^{N}\pi_i^{-1}I_i\delta_i\boldsymbol{h}_{ni}(\boldsymbol{\theta})+o_p(n^{-0.5}) \\ &= N^{-1}\sum_{i=1}^{N}\delta_i\boldsymbol{h}_{ni}(\boldsymbol{\theta})+O_p(n^{-0.5}) \\ &= o_p(1). \end{aligned} \tag{95}$$

Then $||\boldsymbol{G}_n(\boldsymbol{\theta})-\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})||\to 0$. By the law of large numbers $|\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{\theta})|\to 0$. Condition 1 of Pakes and Pollard (1989) holds by the triangle inequality.

To demonstrate condition 2, note that

$$\begin{aligned} \boldsymbol{G}_n(\boldsymbol{\theta})-\boldsymbol{G}(\boldsymbol{\theta})-\boldsymbol{G}_n(\boldsymbol{\theta}_o) &= \boldsymbol{G}_n(\boldsymbol{\theta})-\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})+\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{\theta}) \\ &\quad -(\boldsymbol{G}_n(\boldsymbol{\theta}_o)-\boldsymbol{G}_N(\boldsymbol{\theta}_o,\boldsymbol{y})+\boldsymbol{G}_N(\boldsymbol{\theta}_o,\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{\theta}_o)). \end{aligned} \tag{96}$$

By the mean value theorem (following the argument of Chen and Yu (2016)),

$$\begin{aligned} \boldsymbol{G}_n(\boldsymbol{\theta})-\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})-(\boldsymbol{G}_n(\boldsymbol{\theta}_o)-\boldsymbol{G}_N(\boldsymbol{\theta}_o,\boldsymbol{y})) &= \frac{1}{N}\sum_{i=1}^{N}[(\pi_i^{-1}I_i-1)\dot{\boldsymbol{g}}_\theta(y_i,x_i,\boldsymbol{\theta})](\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}_o) \\ &\quad +B_{2,HT}(\boldsymbol{\theta})-B_{2,HT}(\boldsymbol{\theta}_o)+B_{3,HT}(\boldsymbol{\theta})-B_{3,HT}(\boldsymbol{\theta}_o). \end{aligned}$$

By (21) and by the assumption that $||(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}_o)|\leq|(\boldsymbol{\theta}-\boldsymbol{\theta}_o)||\to 0$,

$$\frac{1}{N}(\sum_{i=1}^{N}[(\pi_i^{-1}I_i-1)\dot{\boldsymbol{g}}_{i,\theta}(y_i;\boldsymbol{\theta})](\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}_o)=o_p(n^{-0.5}), \tag{97}$$

where $\dot{\boldsymbol{g}}_{i,\theta}(y_i;\boldsymbol{\theta})$ is the matrix of derivatives of $\boldsymbol{g}_i(y_i;\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. Similarly,

$$\begin{aligned} B_{2,HT}(\boldsymbol{\theta}) - B_{2,HT}(\boldsymbol{\theta}_o) &= N^{-1}\sum_{i=1}^{N}(1-\delta_i)(\pi_i^{-1}I_i - 1)(\dot{\boldsymbol{\mu}}_{g|x,\theta}(x_i,\boldsymbol{\theta}) - \dot{\boldsymbol{g}}_{i,\theta}(y_i,\boldsymbol{\theta}))(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) \\ &= o_p(n^{-0.5}) \end{aligned}$$

because the assumption that the derivatives have bounded $2+\delta$ moments ensures (21) and because $||(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_o)|| \leq ||(\boldsymbol{\theta} - \boldsymbol{\theta}_o)|| \to 0$. By a similar argument, and assuming second derivatives have bounded 2+δ moments, $B_{3,HT}(\boldsymbol{\theta}) - B_{3,HT}(\boldsymbol{\theta}_o) = o_p(n^{-0.5})$. By the assumption that the derivatives of $\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y})$ and $\boldsymbol{G}(\boldsymbol{\theta})$ are bounded and have bounded moments, $\sqrt{N}(\boldsymbol{G}_N(\boldsymbol{\theta}) - \boldsymbol{G}(\boldsymbol{\theta})) \xrightarrow{d} N(0,\boldsymbol{\Sigma}_{gg})$, where $\boldsymbol{\Sigma}_{gg} = V\{\boldsymbol{g}_i(y_i;\boldsymbol{\theta})\}$. Therefore, an application of the mean value theorem gives,

$$\boldsymbol{G}_N(\boldsymbol{\theta},\boldsymbol{y}) - \boldsymbol{G}(\boldsymbol{\theta}) - (\boldsymbol{G}_N(\boldsymbol{\theta}_o,\boldsymbol{y}) - G(\boldsymbol{\theta}_o)) = o_p(n^{-0.5}). \tag{98}$$

Condition 2 of Pakes and Pollard (1989) holds by the triangle inequality.

C: Computation of Weights for CEAP

The CEAP sample can be approximated as an unequal probability with-replacement sample. The selection probabilities reflect the two-phase process of selection into the NRI core panel, 2002 rotation panel, or 2003 rotation panel. The use of the approximation that the sample is a with-replacement sample is justified because the sampling rate for the foundation sample is relatively small (typically 2%-6%).

The first phase selection probability is based on the stratification procedure defining the NRI foundation sample. Let $p_{f,k}$ denote the probability that segment k is selected into the NRI foundation sample. In a typical stratum defining the NRI foundation sample, 2 segments out of 48 are selected, leading to a foundation probability of 4%.

As discussed above, the sample for the 2003-2005 CEAP survey was approximately the core and 2002-2003 rotation panels. The design for selecting the core and 2002-2003 rotation panels is approximately a stratified sample, with sample classes as strata. Denote the second-phase sample consisting of the core and 2002-2003 rotations as S_2, and let $h(k)$ be the sample class containing segment k. An approximation for the probability of selecting segment k into S_2 is

$$p_{2k} = n_{h(k)} N_{h(k)}^{-1}, \tag{99}$$

where $N_{h(k)}$ and $n_{h(k)}$ are the number of segments in the NRI foundation and S_2, respectively, in sample class h containing segment k. An initial weight reflecting the two-phase process of selecting the sample S_2 is given by,

$$W_{k,1} = p_{f,k}^{-1} D_k p_{2k}^{-1}. \tag{100}$$

A benchmarking procedure is used to define a final weight. Define the estimated area for sample class h based on the foundation segments by,

$$\hat{A}_h = \sum_{\{k:h(k)=h\}} p_{f,k}^{-1} D_k, \tag{101}$$

where D_k is the area of segment k. The weight $W_{k,1}$ is benchmarked to the estimated area $\hat{A}_h$ through a ratio adjustment to construct the final weight $W_{k,2}$, defined,

$$W_{k,2} = W_{k,1} \hat{A}_h \left(\sum_{\{k:h(k)=h\}} W_{k,1} \right)^{-1}. \tag{102}$$

We approximate the CEAP sample as an unequal probability with replacement sample, with selection probability defined by,

$$\pi_k = W_{k,2}^{-1} D_k. \tag{103}$$

One point per segment is selected for the CEAP sample. The selection probability for point i in segment k is defined,

$$\pi_{ik} = \pi_k m_{k,e}^{-1}, \tag{104}$$

where $m_{k,e}$ is the number of cropland points designated as eligible for the 2003-2005 CEAP survey in segment k. The corresponding point weight, denoted W_{ik}, is defined by,

$$W_{ik} = \pi_{ik}^{-1}. \tag{105}$$

The second-order selection probabilities are obtained from the π_k using properties of with-replacement sampling designs. Under the assumption that selecting the same segment twice is negligible, the probability of selecting segment k on a single draw is defined,

$$p_k = 1 - (1 - \pi_k)^{1/n}, \tag{106}$$

where n, as defined above, is the number of segments in the CEAP sample. The joint inclusion probability for segments k and j ($j \neq k$) is then given by,

$$\pi_{jk} = 1 - [(1 - p_j)^n + (1 - p_k)^n - (1 - p_i - p_k)^n] \tag{107}$$

(Sarndal, Swenson, and Wretman, 2003). By the one point per segment rule, the joint inclusion probabilities for point i and point ℓ in segments j and k, respectively, are given by,

$$\begin{aligned} \pi_{j(i)k(\ell)} &= \pi_{jk} m_{j,e}^{-1} m_{k,e}^{-1}, \quad j \neq k \\ &= \pi_i \quad j = k, i = \ell \\ &= 0 \text{ otherwise.} \end{aligned} \tag{108}$$

D: Additional Details on Estimators in Simulations

This appendix presents details of the estimators in the simulations. The function $\boldsymbol{g}_i(y_i; \boldsymbol{\theta})$ defining the estimator of $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ is,

$$\boldsymbol{g}_i(y_i; \boldsymbol{\theta}) = \begin{pmatrix} (y_i - \theta_1) \\ (y_i - \theta_1)^2 - \theta_2 \\ (y_i - \theta_1)(x_i - \mu_x) - \theta_3 \theta_2 \sigma_x^2 \end{pmatrix}, \tag{109}$$

where $(\mu_x, \sigma_x^2) = (E[x_i], V\{x_i\})$. The resulting estimators are

$$\hat{\theta}_1 = \left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} \tilde{y}_i \tag{110}$$

$$\hat{\theta}_2 = \left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} (\tilde{y}_i - \hat{\theta}_1)^2,$$

and

$$\hat{\theta}_3 = \frac{\left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} (\tilde{y}_i - \hat{\theta}_1)(x_i - \hat{\mu}_x)}{\sqrt{[\left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} (\tilde{y}_i - \hat{\theta}_1)^2][\left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} (x_i - \hat{\mu}_x)^2]}}, \tag{111}$$

where $\tilde{y}_i = \delta_i y_i + (1 - \delta_i) J^{-1} \sum_{j=1}^{J} \boldsymbol{B}(x_i) \hat{\boldsymbol{\beta}}_{\tau_j}$, and

$$\hat{\mu}_x = (\sum_{i=1}^{n} \pi_i^{-1})^{-1} (\sum_{i=1}^{n} \pi_i^{-1} x_i). \tag{112}$$

The estimator $\hat{\theta}_4$ of θ_4 is a ratio estimator. The numerator is an estimator of $E[y_i I[x_i \le 0.65]]$, and the denominator is an estimator of $E[I[x_i \le 0.65]]$. The g_i defining the numerator is,

$$g_{4,i}(y_i; \theta_{4,n}) = (y_i I[x_i \le 0.65] - \theta_{4,n}) \tag{113}$$

where $\theta_{4,n} = E[y_i I[x_i \le 0.65]]$. The estimator of $\hat{\theta}_4$ is then defined,

$$\hat{\theta}_4 = \frac{\left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} y_i I[x_i < 0.65]}{\left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} I[x_i < 0.65]}. \tag{114}$$

The estimator of the variance of $\hat{\theta}_4$ is obtained by a Taylor approximation for a ratio. The estimator of θ_5 does not fall in the framework of Section 3 because $I[a \le 8]$ is a non-smooth function of a; however, we evaluate the empirical properties of estimator $\hat{\theta}_2$ defined,

$$\hat{\theta}_5 = \left(\sum_{i=1}^{n} \pi_i^{-1}\right)^{-1} \sum_{i=1}^{n} \pi_i^{-1} \left\{ \delta_i I[y_i \le 8] + (1 - \delta_i) J^{-1} \sum_{j=1}^{J} I[y_{ij} \le 8] \right\}. \tag{115}$$

E: Discussion Expanded

Choice of τ_j

One difference between the method that we present and the Chen and Yu (2016) method involves the determination of τ_j. In Chen and Yu (2016), τ_j is randomly generated from a uniform distribution on the interval $(0, 1)$. The use of random τ_j can lead to poor variance estimates if generated values of τ_j are close to 0 or 1. We chose to use the midpoint approximation instead because this led to improved variance estimates with little impact on the bias of the estimator. In the simulation, we use $J = 100$. Simulations not presented here indicate that a choice of $J = 50$ is also reasonable.

Improvements to estimation of the quantile curves

In the simulations and application, simple methods are used to select the tuning parameter and number of knots for estimation of the quantile curves. Development of automated methods to select these nuisance parameters, appropriate for selection of multiple quantiles in a complex survey setting is an area for future research. Similarly, estimation of the quantile curves subject to a restriction that the estimated curves are non-overlapping has potential to improve estimation of the derivatives needed for the variance estimator.

Choice of b_i

An interesting issue involves the choice of b_i. A $b_i = w_i^{-1}$ leads to an unweighted estimator, while $b_i = 1$ incorporates weights. For the simulations and data analysis presented here, we use $b_i = 1$. A comparison of estimators with $b_i = 1$ to $b_i = w_i^{-1}$ indicates that the unweighted and weighted QRI estimators have similar properties for a range of designs. An investigation of conditions under which an estimator with $b_i = w_i^{-1}$ is nearly unbiased would be an interesting and challenging area for future work.

References

Chen, S. and Yu, C. (2016). Parameter Estimation Through Semiparametric Quantile Regression Imputation. *Electronical Journal of Statistics*, 10, 3621–3647.

Knight, K. (1998). Limiting Distributions for L1 Regression Estimators Under General Conditions. *The Annals of Statistics*, 26, 755–770.

Pakes, A. and Pollard D. (1989). Simulation and the Asymptotic of Optimization Estimators, *Econometrica* 57(4),1027–1057.

Sarndal, C.E., Swensson, B., and Wretman, J. (2003). *Model Assisted Survey Sampling*, Springer Science & Business Media.