Estimation for nonignorable missing response or covariate using semi-parametric quantile regression imputation and a parametric propensity score model.

Department of Statistics, Iowa State University

Supplementary Material

In this supplement, we provide additional explanation omitted from the main manuscript for brevity. In Section S1, we outline proofs of Theorem 1 and Theorem 2 given in Section 3 of the main manuscript. In Section S2, we define explicit expressions for the plug-in variance estimators. In Section S3, we expand on the propensity score adjustment introduced in Section 3.2 of the main manuscript. Finally, in Section S4, we outline an identification condition for a multivariate covariate.

S1 Proofs

S1.1 Corollary to Lemma 1 and Proof

Corollary 1. *Under the assumptions of Lemma 1,*

$$\boldsymbol{B}(x)'\hat{\boldsymbol{\beta}}_{y|x}(\tau) - b_\tau^\lambda(x) = O_p\left(\frac{K_{n_1,y}}{n_1}\right).$$

Proof. By Chen and Yu (2016) and Yoshida (2013), $b_\tau^\lambda(x_i) = O_p(K_{n_1,y}^{-(p_{y|x}+1)})$, and $b_\tau^a(x_i) = O_p(K_{n_1,y}^{-(p_{y|x}+1)})$. By assumption, $K_{n_1,y} = O(n_1^{1/(2p_{y|x}+3)})$. Then,

$$O(K_{n_1,y}^{-(p_{y|x}+1)} n_1^{0.5} K_{n_1,y}^{-0.5}) = O\left(\frac{\sqrt{n_1}}{n_1^{\frac{2(p_{y|x}+1)}{2(2p_{y|x}+3)}}}\left(\frac{1}{n_1^{\frac{1}{2(2p_{y|x}+3)}}}\right)\right) = O(1).$$

□

S1.2 Proof of Theorem 1

The three additional regularity conditions below are required for the approximation with finite J:

1. $f(y \mid x, \delta = 1)$ is three times differentiable and bounded on $[M_{1y}, M_{2y}]$.

2. $\dot{f}(y \mid x, \delta = 1) = \partial f(t \mid x, \delta = 1)/\partial t|_{t=y}$ is continuous and bounded on $[M_{1y}, M_{2y}]$ for any $x \in [M_{1x}, M_{2x}]$

3. $\ddot{f}(y \mid x, \delta = 1) = \partial^2 f(t \mid x, \delta = 1)/\partial t^2|_{t=y}$ is continuous and bounded on $[M_{1y}, M_{2y}]$ for any $x \in [M_{1x}, M_{2x}]$.

Proof of Theorem 1. : We obtain a linear approximation for $\hat{\boldsymbol{\phi}}_2 = (\hat{\phi}_{02}, \hat{\phi}_{12}, \hat{\phi}_{22})'$.

As notation, let

$$\boldsymbol{S}(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_y) = \frac{1}{n}\sum_{i \in A_{12}} (\delta_{1i} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i} := \frac{1}{n}\sum_{i \in A_{12}} \boldsymbol{S}_i(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_{yi}), \quad \text{(S1.1)}$$

and $\boldsymbol{q}_y = \{\boldsymbol{q}_{yi} : i \in A_{12}\}$. The estimator of $\boldsymbol{\phi}_2$ satisfies $\boldsymbol{S}(\hat{\boldsymbol{\phi}}_2 \mid \hat{\boldsymbol{q}}_y) = \boldsymbol{0}$, where $\boldsymbol{S}(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_y)$ is defined in (S1.1). In the derivation of the score equation (S1.1), we use the fact that

$$\frac{\partial \hat{h}_y(-\phi_{22}, x_i)}{\partial \phi_{22}} = -E_{2,J}(y_i \mid x_i).$$

Then,

$$\begin{aligned}
0 = \boldsymbol{S}(\hat{\boldsymbol{\phi}}_2 \mid \hat{\boldsymbol{q}}_y) &= \boldsymbol{S}_\infty(\boldsymbol{\phi}_2) + [\boldsymbol{S}(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_y) - \boldsymbol{S}_\infty(\boldsymbol{\phi}_2)] \\
&\quad - \boldsymbol{I}_{n,\phi_2}(\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2) + \frac{1}{n}\sum_{i=1}^n (\delta_{1i} + \delta_{2i}) \sum_{j=1}^J \boldsymbol{d}_{qij}(\hat{q}_{\tau_j}(x_i) - q_{\tau_j}(x_i)) \\
&\quad + \boldsymbol{R}_{n,\phi\phi} + \boldsymbol{R}_{n,\phi q} + \boldsymbol{R}_{n,qq},
\end{aligned}$$

where $\boldsymbol{S}_\infty(\boldsymbol{\phi}_2) = n^{-1}\sum_{i \in A_{12}} \boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2)$,

$$\begin{aligned}
\boldsymbol{R}_{n,\phi\phi} &= \frac{1}{n}\sum_{i=1}^n (\delta_{1i} + \delta_{2i}) \frac{1}{2}\sum_{k=0}^2 \sum_{k'=0}^2 \boldsymbol{W}^*_{\phi\phi kk',i}(\hat{\phi}_{k2} - \phi_{k2})(\hat{\phi}_{k'2} - \phi_{k'2}) \\
&\quad + \frac{1}{n}\sum_{i=1}^n (\delta_{1i} + \delta_{2i}) \frac{1}{2}\sum_{k=0}^2 \sum_{k'=0}^2 (\check{\boldsymbol{W}}^*_{\phi\phi kk',i} - \boldsymbol{W}^*_{\phi\phi kk',i})(\hat{\phi}_{k2} - \phi_{k2})(\hat{\phi}_{k'2} - \phi_{k'2}) - (\check{\boldsymbol{I}}_{n,\boldsymbol{\phi}_2} - \boldsymbol{I}_{n,\boldsymbol{\phi}_2})(\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2) \\
\boldsymbol{R}_{n,\phi q} &= \frac{1}{n}\sum_{i=1}^n (\delta_{1i} + \delta_{2i}) \frac{1}{2}\sum_{k=0}^2 \sum_{j=1}^J \boldsymbol{W}^*_{\phi qkj,i}(\hat{\phi}_{k2} - \phi_{k2})(\hat{q}_{\tau_j}(x_i) - q_{\tau_j}(x_i)) \\
&\quad + \frac{1}{n}\sum_{i=1}^n (\delta_{1i} + \delta_{2i}) \frac{1}{2}\sum_{k=0}^2 \sum_{j=1}^J (\check{\boldsymbol{W}}^*_{\phi qkj,i} - \boldsymbol{W}^*_{\phi qkj,i})(\hat{\phi}_{k2} - \phi_{k2})(\hat{q}_{\tau_j}(x_i) - q_{\tau_j}(x_i)),
\end{aligned}$$

$$\boldsymbol{R}_{n,qq} = \frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\frac{1}{2}\{\sum_{j=1}^{J}\boldsymbol{W}^*_{qqjj,i}(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))^2$$

$$+\sum_{j=1}^{J}\sum_{j'\neq j}\boldsymbol{W}_{qqjj',i}(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))(\hat{q}_{\tau_{j'}}(x_i)-q_{\tau_{j'}}(x_i))\}$$

$$+\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\frac{1}{2}\sum_{j=1}^{J}(\check{\boldsymbol{W}}^*_{qqjj,i}-\boldsymbol{W}^*_{qqjj,i})(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))^2$$

$$+\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\frac{1}{2}\sum_{j=1}^{J}\sum_{j'\neq j}(\check{\boldsymbol{W}}^*_{qqjj',i}-\boldsymbol{W}^*_{qqjj',i})(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))(\hat{q}_{\tau_{j'}}(x_i)-q_{\tau_{j'}}(x_i))$$

$$+\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i)),$$

$$\boldsymbol{W}^*_{\phi\phi kk',i} = E[\frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial\phi_{2k}\partial\phi_{2k'}}\mid x_i, i\in A_{12}]|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*},\quad \check{\boldsymbol{W}}^*_{\phi\phi kk',i} = \frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial\phi_{2k}\partial\phi_{2k'}}|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*}\quad k,k'=0,1,2,$$

$$\boldsymbol{W}^*_{\phi qkj,i} = E[\frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial\phi_{2k}\partial q_{\tau_j}(x_i)}\mid x_i, i\in A_{12}]|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*},\quad \check{\boldsymbol{W}}^*_{\phi qkj,i} = \frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial\phi_{2k}\partial q_{\tau_j}(x_i)}|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*},\quad k=0,1,2, j=1,\ldots,J,$$

$$\boldsymbol{W}^*_{qqjj',i} = E[\frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial q_{\tau_j}(x_i)\partial q_{\tau_{j'}}(x_i)}\mid x_i, i\in A_{12}]|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*},\quad \check{\boldsymbol{W}}^*_{qqjj',i} = \frac{\partial^2\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial q_{\tau_j}(x_i)\partial q_{\tau_{j'}}(x_i)}|_{\boldsymbol{\phi}_2^*,\boldsymbol{q}^*},\quad j,j'=1,\ldots,J$$

$$\check{\boldsymbol{d}}_{qij} = \frac{\partial\boldsymbol{S}_i(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})}{\partial q_{\tau_j}(x_i)},\quad \boldsymbol{d}_{qij} = \pi_{12i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi}))w_{2ij}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})\phi_{22}\boldsymbol{z}_{2i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi}),$$

$\boldsymbol{q}^* = \{q_{\tau_j}(x_i)^* : j = 1,\ldots,J;\delta_i = 1\}$, $\boldsymbol{\phi}_2^* = (\phi_{20}^*,\phi_{21}^*,\phi_{22}^*)$, $q_{\tau_j}(x_i)^*$ is between $\hat{q}_{\tau_j}(x_i)$ and $q_{\tau_j}(x_i)$, and ϕ_{k2}^* is between $\hat{\phi}_{k2}$ and ϕ_{k2} for $k=0,1,2$.

We first consider $\boldsymbol{S}(\boldsymbol{\phi}_2\mid\boldsymbol{q}_y)-\boldsymbol{S}_\infty(\boldsymbol{\phi}_2)$. We expand this difference as

$$\boldsymbol{S}(\boldsymbol{\phi}_2\mid\boldsymbol{q}_y)-\boldsymbol{S}_\infty(\boldsymbol{\phi}_2) = \frac{1}{n}\sum_{i\in A_{12}}(\delta_{1i}-\pi_{12i\infty}(\boldsymbol{\phi}_2))(\boldsymbol{z}_{2i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})-\boldsymbol{z}_{2i\infty}(\boldsymbol{\phi}_2))$$

$$+\frac{1}{n}\sum_{i\in A_{12}}(\pi_{12i\infty}(\boldsymbol{\phi}_2)-\pi_{12i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi}))\boldsymbol{z}_{2i\infty}(\boldsymbol{\phi}_2)$$

$$+\frac{1}{n}\sum_{i\in A_{12}}(\pi_{12i\infty}(\boldsymbol{\phi}_2)-\pi_{12i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi}))(\boldsymbol{z}_{2i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})-\boldsymbol{z}_{2i\infty}(\boldsymbol{\phi}_2)).$$

We next express $\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})$ as a function of τ_1 alone as

$$\pi_{12i}(\boldsymbol{\phi}_2, \tau_1) = \frac{\exp(-\phi_{20} - \phi_{21}x_i + \hat{h}_{yi}(\tau_1))}{1 + \exp(-\phi_{20} - \phi_{21}x_i + \hat{h}_{yi}(\tau_1))},$$

where $\hat{h}_{yi}(\tau_1) = -\log(\mu_{1i}(\tau_1))$,

$$\mu_{1i}(\tau_1) = J^{-1}\sum_{j=1}^{J} \exp(\phi_{22}q_i(\tau_1 + (j-1)/J)),$$

Let $\mu_{1i\infty} = \int_0^1 \exp(\phi_{22}q_i(\tau))d\tau$. We now use Theorem 5.4.3 of Fuller (2005) to verify that

$$\mu_{1i}(\tau_1) = \mu_{1i}(1/(2J)) + O(J^{-1}), \tag{S1.2}$$

for any x_i. First, $E[|\tau_1 - 0.5/J|^2] = 1/(12J^2)$. By assumption, $\partial f(F^{-1}_{y|x,\delta=1}(\tau) \mid x, \delta = 1)/\partial\tau$ is continuous on $(0,1)$. Therefore, $\partial f(F^{-1}_{y|x,\delta=1}(\tau) \mid x, \delta = 1)/\partial\tau$ and bounded on a closed interval containing $1/(2J)$. Then, the assumptions of Theorem 5.4.3 of Fuller (2005) with $\delta = 1$, $\alpha = 2$, and $s = 1$ are satisfied, and (S1.2) holds for any x_i. Observe that $\mu_{1i}(1/(2J))$ is mid-point approximation for $\mu_{1i\infty}$. By assumption, $\partial^2 f(F^{-1}_{y|x,\delta=1}(\tau) \mid x, \delta = 1)/\partial\tau^2$ is continuous and bounded over $[0,1]$. Then, $\mu_{1i}(1/(2J)) = \mu_{1i\infty} + O(J^{-1})$.

We then express $\pi_{12i}(\boldsymbol{\phi}_2, \tau_1)$ as

$$\pi_{12i}(\boldsymbol{\phi}_2, \mu_{1i}(\tau_1)) = \frac{\exp(-\phi_{20} - \phi_{21}x_i + \hat{h}_{yi}(\mu_{1i}(\tau_1)))}{1 + \exp(-\phi_{20} - \phi_{21}x_i + \hat{h}_{yi}(\mu_{1i}(\tau_1))}. \tag{S1.3}$$

A first-order Taylor approximation then gives

$$\pi_{12i}(\boldsymbol{\phi}_2, \mu_{1i}(\tau_1)) - \pi_{12i}(\boldsymbol{\phi}_2, \mu_{1i\infty}) = \pi_{12i}(\boldsymbol{\phi}_2, \tilde{\mu}_{1i\infty})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \tilde{\mu}_{1i\infty}))(-\tilde{\mu}^{-1}_{1i\infty})(\mu_{1i}(\tau_1) - \mu_{1i\infty}) = O(J^{-1}),$$

where $\tilde{\mu}_{1i}$ is between $\mu_{1i}(\tau_1)$ and $\mu_{1i\infty}$. We can similarly express $z_{1i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})$ as a function of τ_1 as

$$z_{1i}(\boldsymbol{\phi}_2, \tau_1) = \frac{\sum_{j=1}^{J} q_i(\tau_1 + j/(J-1))\exp(\phi_{22}) q_i(\tau_1 + (j-1)/J)}{\sum_{j=1}^{J} \exp(\phi_{22}) q_i(\tau_1 + (j-1)/J)}.$$

Because $z_{1i}(\boldsymbol{\phi}_2, \tau_1)$ is a continuous function of τ_1, $z_{1i}(\boldsymbol{\phi}_2, \tau_1) = z_{1i\infty}(\boldsymbol{\phi}_2) + o_p(1)$. Then, $\boldsymbol{S}(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_y) - \boldsymbol{S}_\infty(\boldsymbol{\phi}_2) = O_p(J^{-1}) = o_p(n^{-0.5})$, where the second equality follows from our assumption about the order of J. Direct differentiation (see Section S1.3 of this supplement) shows that $\boldsymbol{W}_{\phi qkj,i} = O_p(J^{-1})$, $\boldsymbol{W}_{qqjj,i} = O_p(J^{-1})$, $\boldsymbol{W}_{qqjj',i} = O_p(J^{-2})$, $\check{\boldsymbol{W}}_{\phi qkj,i} = O_p(J^{-1})$, $\check{\boldsymbol{W}}_{qqjj,i} = O_p(J^{-1})$, and $\check{\boldsymbol{W}}_{qqjj',i} = O_p(J^{-2})$. To obtain the order of $\boldsymbol{R}_{n,qq}$, we expand the last term in the sum defining $\boldsymbol{R}_{n,qq}$ as

$$\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))$$
$$= \frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})\times$$
$$\times\left\{\boldsymbol{B}(x_i)'\frac{1}{n_1}\sum_{k=1}^{n}\delta_{1k}\boldsymbol{H}_{n_1,y|x}^{-1}(\tau_j)\boldsymbol{B}(x_k)\psi_{\tau_j}(e_{y|x,k}(\tau_j)) + b_{\tau_j}^{a}(x_i) + b_{\tau_j}^{\lambda}(x_i) + o_p\left(\sqrt{\frac{K_{n_1,y}}{n_1}}\right)\right\}.$$

By the form of $\check{\boldsymbol{d}}_{qij}$,

$$|n^{-1}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})| \le |\frac{1}{n}\sum_{i\in A_{12}}(\delta_{1i}-\pi_{12i\infty}(\boldsymbol{\phi}_2))|O(J^{-1})$$
$$+|\frac{1}{n}\sum_{i\in A_{12}}(\pi_{12i}(\boldsymbol{\phi}_2,\boldsymbol{q}_{yi})-\pi_{12i\infty}(\boldsymbol{\phi}_2))|O(J^{-1}) = O_p(J^{-1}n^{-0.5}).$$

Then, $n^{-1}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})\boldsymbol{B}(x_i)'[n_1^{-1}\sum_{k=1}^{n}\delta_{1k}\boldsymbol{H}_{n_1,y|x}^{-1}(\tau_j)\boldsymbol{B}(x_k)\psi_{\tau_j}(e_{y|x,k}(\tau_j))] = O_p(n^{-1})$, $n^{-1}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})(b_{\tau_j}^{a}(x_i)+b_{\tau_j}^{\lambda}(x_i)) = O_p(n^{-0.5}K_{n_1,y}^{-(p+1)})$, and $n^{-1}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}(\check{\boldsymbol{d}}_{qij}-\boldsymbol{d}_{qij})o_p\left(\sqrt{K_{n_1,y}n_1^{-1}}\right) = O_p(n^{-0.5})o_p\left(\sqrt{K_{n_1,y}n_1^{-1}}\right) =$ $o_p(n^{-0.5})$. Therefore, $\boldsymbol{R}_{n,qq} = o_p(n^{-0.5})$. Then,

$$(\hat{\boldsymbol{\phi}}_2-\boldsymbol{\phi}_2)-\boldsymbol{I}_{n,\phi_2}^{-1}\boldsymbol{R}_{n,\phi\phi}-\boldsymbol{I}_{n,\phi_2}^{-1}\boldsymbol{R}_{n,\phi q} = \boldsymbol{I}_{n,\phi_2}^{-1}\left[\boldsymbol{S}_\infty(\boldsymbol{\phi}_2)+\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}\boldsymbol{d}_{qij}(\hat{q}_{\tau_j}(x_i)-q_{\tau_j}(x_i))\right] \tag{S1.4}$$

$$\begin{aligned}&+o_p(n^{-0.5})\\ &= \boldsymbol{I}_{n,\phi_2}^{-1}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2)\\ &+\boldsymbol{I}_{n,\phi_2}^{-1}\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}\boldsymbol{d}_{qij}\boldsymbol{B}(x_i)'\left(\frac{1}{n_1}\sum_{k=1}^{n}\delta_{1k}\boldsymbol{H}_{n_1,y|x}^{-1}(\tau_j)\boldsymbol{B}(x_k)\psi_{\tau_j}(e_{y|x,k}(\tau_j))\right)\\ &+\boldsymbol{I}_{n,\phi_2}^{-1}\frac{1}{n}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})\sum_{j=1}^{J}\boldsymbol{d}_{qij}[b_{\tau_j}^{\lambda}(x_i)+b_{\tau_j}^{a}(x_i)]+o_p(n^{-0.5}).\end{aligned}$$

Let $K_n = K_{n_1,y}$. By definition of $b_{\tau_j}^a(x)$,

$$E[b_{\tau_j}^a(X)^2 \mid \delta_1+\delta_2=1] = \sum_{k=1}^{K_n}\int_{\kappa_{k-1}}^{\kappa_k}\frac{q^{(p+1)}(x)^2}{K_n^{2p+2}[(p+1)!]^2}B_{rp}((x-\kappa_{k-1})/K_n^{-1})^2dF(x\mid\delta_1+\delta_2=1) \tag{S1.5}$$

$$\begin{aligned}&= \sum_{k=1}^{K_n}\int_{\kappa_{k-1}}^{K_n(\kappa_k-\kappa_{k-1})}\frac{q^{(p+1)}(K_n^{-1}u+\kappa_{k-1})^2}{K_n^{2p+3}[(p+1)!]^2}B_{rp}(u)^2dF(u\mid\delta_1+\delta_2=1)\\ &= O(K_n^{-(2p+2)}),\end{aligned}$$

for $u=(x-\kappa_{k-1})K_n$, where $B_{rp}(u)^2$ and $q^{(p+1)}(x)$ are bounded. As in Chen and Yu (2016), $E[(b_\tau^\lambda(X))^2] = O(K_n^{-(2p+2)})$ (see below C.4 of the supplement of Chen and Yu (2016)). Note that $\{X_i : \delta_{1i}+\delta_{2i}=1\}$ are *iid*. Therefore, $V\{n^{-1}\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})b_{\tau_j}^a(X_i)\} = O(n^{-1}K_n^{-2(p+1)})$, and $V\{n^{-1}\sum_{i=1}^{n}(\delta_{1i}+$

$\delta_{2i})b^{\lambda}_{\tau_j}(X_i)\} = O(n^{-1}K_n^{-2(p+1)})$. Also, $n^{-1}\sum_{i=1}^n \boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) = O_p(n^{-0.5})$,

$$n_1^{-1}\sum_{i=1}^n \delta_{1i}\boldsymbol{H}^{-1}_{n_1,y|x}(\tau_j)\boldsymbol{B}(x_i)\psi_{\tau_j}(e_{y|x,i}(\tau_j)) = O_p(\sqrt{K_n/n})$$

and $E[\boldsymbol{B}(x) \mid \delta_1 + \delta_2 = 1] = O(K_n^{-0.5})$, as in Fact 1 from Chen and Yu (2016, supplement). Then, the right hand side of (S1.4) is $O_p(n^{-0.5})$. It follows that $\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2 = O_p(n^{-0.5})$, $\boldsymbol{R}_{n,\phi\phi} = o_p(n^{-0.5})$, and $\boldsymbol{R}_{n,\phi q} = o_p(n^{-0.5})$. Define $\boldsymbol{\ell}_{kj} = \boldsymbol{H}^{-1}_{n_1,y|x}(\tau_j)\boldsymbol{B}(x_k)\psi_{\tau_j}(e_{y|x,k}(\tau_j))$. From (S1.4),

$$\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2 = \boldsymbol{I}^{-1}_{n,\phi_2}\left\{\frac{1}{n}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\left(\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) + \sum_{j=1}^J \boldsymbol{d}_{qij}\left[\boldsymbol{B}(x_i)'\frac{1}{n_1}\sum_{k=1}^n \delta_{1k}\boldsymbol{\ell}_{kj} + b^a_{\tau_j}(x_i) + b^{\lambda}_{\tau_j}(x_i)\right]\right)\right\}$$
$$+ o_p(\sqrt{\frac{1}{n}})$$
$$= \boldsymbol{I}^{-1}_{n,\phi_2}\left\{\frac{1}{n}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) + \sum_{j=1}^J[\frac{1}{n}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\boldsymbol{d}_{qij}\boldsymbol{B}(x_i)']\frac{1}{n_1}\sum_{k=1}^n\delta_{1k}\boldsymbol{\ell}_{kj}\right\}$$
$$+ \boldsymbol{I}^{-1}_{n,\phi_2}\sum_{j=1}^J\frac{1}{n}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\boldsymbol{d}_{qij}[b^a_{\tau_j}(x_i) + b^{\lambda}_{\tau_j}(x_i)] + o_p(\sqrt{\frac{1}{n}})$$
$$= \boldsymbol{I}^{-1}_{n,\phi_2}\frac{1}{n}\sum_{i=1}^n\left\{(\delta_{1i}+\delta_{2i})\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) + \sum_{j=1}^J[\frac{1}{n_1}\sum_{i=1}^n(\delta_{1i}+\delta_{2i})\boldsymbol{d}_{qij}\boldsymbol{B}(x_i)']\delta_{1i}\boldsymbol{\ell}_{ij}\right\} + o_p\left(\sqrt{\frac{1}{n}}\right)$$

$$= \boldsymbol{I}^{-1}_{n,\phi_2}\frac{1}{n}\sum_{i=1}^n\{(\delta_{1i}+\delta_{2i})\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2)$$
$$+ \int_{M_{1x}}^{M_{2x}} p_1^{-1}\phi_{22}\pi_{12\infty}(x)(1-\pi_{12\infty}(x))\boldsymbol{z}_{2\infty}(x)\boldsymbol{B}(x)'\sum_{j=1}^J\frac{\exp(\phi_{22}q_{\tau_j}(x))\delta_{1i}\boldsymbol{\ell}_{ij}}{\sum_{j=1}^J\exp(\phi_{22}q_{\tau_j}(x))}dF(x \mid \delta_1+\delta_2=1)\}$$
$$+ o_p(n^{-0.5}),$$

$$= \boldsymbol{I}_{n,\phi_2}^{-1} \frac{1}{n} \sum_{i=1}^{n} \{(\delta_{1i} + \delta_{2i}) \boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2)$$

$$+ \delta_{1i} \int_{M_{1x}}^{M_{2x}} p_1^{-1} \phi_{22} \pi_{12\infty}(x)(1 - \pi_{12\infty}(x)) \boldsymbol{z}_{2\infty}(x) \boldsymbol{B}(x)' \sum_{j=1}^{J} \frac{\exp(\phi_{22} q_{\tau_j}(x)) \boldsymbol{\ell}_{ij}}{\sum_{j=1}^{J} \exp(\phi_{22} q_{\tau_j}(x))} dF(x \mid \delta_1 + \delta_2 = 1)\}$$

$$+ o_p(n^{-0.5}),$$

$$= \boldsymbol{I}_{n,\phi_2}^{-1} \frac{1}{n} \sum_{i=1}^{n} \{(\delta_{1i} + \delta_{2i}) \boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) + \delta_{1i} \boldsymbol{L}_{i\infty} + \delta_{1i} R_{\infty i}\} + o_p(n^{-0.5}),$$

where

$$\boldsymbol{L}_{i\infty} = \phi_{22} \int_{M_{1x}}^{M_{2x}} p_1^{-1} \pi_{12\infty}(x)(1 - \pi_{12\infty}(x)) \boldsymbol{z}_{2\infty}(x) \boldsymbol{B}(x)' \frac{\int_0^1 \exp(\phi_{22} q_\tau(x)) \boldsymbol{\ell}_i(\tau) d\tau}{\int_0^1 \exp(\phi_{22} q_\tau(x)) d\tau} dF(x \mid \delta_1 + \delta_2 = 1)$$

and $R_{\infty i} = \int_{M_{1x}}^{M_{2x}} p_1^{-1} \phi_{22} \pi_{12\infty}(x)(1 - \pi_{12\infty}(x)) \boldsymbol{z}_{2\infty}(x) \boldsymbol{B}(x)' \sum_{j=1}^{J} \frac{\exp(\phi_{22} q_{\tau_j}(x)) \boldsymbol{\ell}_{ij}}{\sum_{j=1}^{J} \exp(\phi_{22} q_{\tau_j}(x))} dF(x \mid \delta_1 + \delta_2 = 1) - \boldsymbol{L}_{i\infty}$. Because $\exp(\phi_{22} q_i(\tau))$ is twice-differentiable in τ, Theorem 5.4.3 of Fuller (2005) and the mid-point approximation for the integral imply that

$$\frac{J^{-1} \sum_{j=1}^{J} \exp(\phi_{22} F_x^{-1}(\tau_1 + (j-1)/J))}{\int_0^1 \exp(\phi_{22} F_x^{-1}(\tau)) d\tau} - 1 = O(J^{-1}),$$

where $F_x^{-1}(\tau) = q_\tau(x)$ is used to emphasize that the quantile is the inverse of the CDF. Because $\boldsymbol{\ell}_i(\tau)$ is not differentiable in τ, Theorem 5.4.3 of Fuller (2005) and the mid-point approximation do not apply. Define

$$\Delta_i = J^{-1} \sum_{j=1}^{J} \exp(\phi_{22} F_x^{-1}(\tau_1 + (j-1)/J)) \boldsymbol{\ell}_i(\tau_1 + (j-1)/J) - \int_0^1 \exp(\phi_{22} F_x^{-1}(\tau)) \boldsymbol{\ell}_i(\tau) d\tau.$$

Then,

$$\Delta_i = J^{-1}\sum_{j=1}^{J}\exp(\phi_{22}F_x^{-1}(\tau_1+(j-1)/J))\boldsymbol{\ell}_i(\tau_1+(j-1)/J) - \sum_{j=1}^{J}\int_{(j-1)/J}^{j/J}\exp(\phi_{22}F_x^{-1}(\tau))\boldsymbol{\ell}_i(\tau)d\tau$$
$$= J^{-1}\sum_{j=1}^{J}C_i(\tau_j)\psi_{\tau_j}(e_{y|x,i}(\tau_j)) - J\sum_{j=1}^{J}\int_0^{1/J}J^{-1}C_i(\tau_j^*)\psi_{\tau_j^*}(e_{y|x,i}(\tau_j^*))d\tau^*,$$

where $\tau_j = \tau_1+(j-1)/J$, $\tau_j^* = \tau^*-(j-1)/J$, and $C_i(\tau) = \exp(\phi_{22}F_x^{-1}(\tau))\boldsymbol{H}_{n_1,y|x}^{-1}(\tau)\boldsymbol{B}(x_i)$.

By the triangle inequality,

$$|\Delta_i| \le |J^{-1}\sum_{j=1}^{J}C_i(\tau_j)\psi_{\tau_j}(e_i(\tau_j)) - J\sum_{j=1}^{J}\int_0^{1/J}J^{-1}C_i(\tau_j)\psi_{\tau_j^*}(e_i(\tau_j^*))d\tau^*|$$
$$+ |J\sum_{j=1}^{J}\int_0^{1/J}\frac{C_i(\tau_j)-C_i(\tau_j^*)}{J}\boldsymbol{B}(x_i)\psi_{\tau_j^*}(e_i(\tau_j^*))d\tau^*|.$$

By the assumption that $\ddot{f}(y \mid x, \delta = 1)$ is continuous $C_i(\tau_j) - C_i(\tau_j^*) = O(J^{-1})$, and

$$|J\sum_{j=1}^{J}\int_0^{1/J}\frac{C_i(\tau_j)-C_i(\tau_j^*)}{J}\boldsymbol{B}(x_i)\psi_{\tau_j^*}(e_i(\tau_j^*))| = O(J^{-1}).$$

Let M be such that $|C_i(\tau)| \le M_i$ for all $\tau \in (0,1)$. Then,

$$|\Delta_i| \le M|J^{-1}\sum_{j=1}^{J}\psi_{\tau_j}(e_{y|x,i}(\tau_j)) - J\sum_{j=1}^{J}\int_0^{1/J}J^{-1}\psi_{\tau_j^*}(e_{y|x,i}(\tau_j^*))d\tau^*|.$$

The indicator function defining $\psi_{\tau_1+(j-1)/J}(e_i(\tau_1+(j-1)/J))$ will change in an interval for exactly one j for every i. For the interval where the indicator changes,

$$|\psi_{\tau_1+(j-1)/J}(e_i(\tau_1+(j-1)/J))/J - \int_0^{1/J}\psi_{\tau_1+(j-1)/J}(e_i(\tau_1+(j-1)/J))| = O(1/J).$$

For all other intervals,

$$|\psi_{\tau_1+(j-1)/J}(e_i(\tau_1+(j-1)/J))/J - \int_0^{1/J} \psi_{\tau_1+(j-1)/J}(e_i(\tau_1+(j-1)/J))| = O(1/J^2).$$

Then, $|R_{\infty i}| = O(J^{-1})$, and

$$\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2 = \boldsymbol{I}_{n,\phi_2}^{-1} \frac{1}{n} \sum_{i=1}^{n} \{(\delta_{1i}+\delta_{2i})\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) + \delta_{1i}\boldsymbol{L}_{i\infty}\} + o_p(n^{-0.5}).$$

By definition of $\psi_\tau(e_{y|x,i}(\tau))$ and exchanging the order of integration, $E\delta_{1i}\boldsymbol{L}_{i\infty} = 0$, and $E\boldsymbol{S}_{i\infty}(\boldsymbol{\phi}_2) - \boldsymbol{0}$. Because Y and X have compact support, second moments exist, and the result follows from the Central Limit Theorem.

□

S1.3 Detailed Derivatives Required for Asymptotic Approximations in Theorem 1

$$\frac{\partial \hat{h}_y(-\phi_{22}, \boldsymbol{q}_{yi})}{\partial \phi_{22}} = -E_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi}), \quad \frac{\partial \hat{h}_y(-\phi_{22}, \boldsymbol{q}_{yi})}{\partial q_{\tau_j}(x_i)} = -\phi_{22} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) \tag{S1.6}$$

$$\frac{\partial}{\partial \phi_{22}} E_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi}) = Var_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi}) \tag{S1.7}$$

$$\frac{\partial}{\partial q_{\tau_j}(x_i)} E_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi}) = w_{2ij} \left(1 + \frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J} \exp(\phi_{22} q_{\tau_j}(x_i))} - \phi_{22} E_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi}) \right)$$

$$\frac{\partial}{\partial \phi_{22}} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) = w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})[q_{\tau_j}(x_i) - E_{2,J}(y_i \mid x_i)] \tag{S1.8}$$

$$\frac{\partial}{\partial q_{\tau_j}(x_i)} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) = w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22} - w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})^2 \phi_{22} \tag{S1.9}$$

$$\frac{\partial}{\partial q_{\tau_{j'}}(x_i)} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) = -w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) w_{2ij'}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22} \tag{S1.10}$$

$$\frac{\partial \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})}{\partial q_{\tau_j}(x_i)} = \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))(-\phi_{22} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) \tag{S1.11}$$

$$\frac{\partial \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})}{\partial \phi_{22}} = \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i} \tag{S1.12}$$

$$\begin{aligned} \boldsymbol{H}_{\phi\phi,i} &:= E\left[\frac{\partial \boldsymbol{S}_i(\boldsymbol{\phi}_2 \mid \boldsymbol{q}_{yi})}{\partial \boldsymbol{\phi}_2} \mid x_i, i \in A_{12} \right] \qquad \text{(S1.13)} \\ &= -\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i}\boldsymbol{z}_{2i}' \end{aligned}$$

$$\check{\boldsymbol{H}}_{\phi\phi,i} = \frac{\partial \boldsymbol{S}_i}{\partial \boldsymbol{\phi}_2} = \boldsymbol{H}_{\phi\phi,i} + (\delta_{1i} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{M}_{ci}, \tag{S1.14}$$

$$\check{\boldsymbol{I}}_{n,\phi_2} = n^{-1} \sum_{i \in A_{12}} \check{\boldsymbol{H}}_{\phi\phi,i}$$

$$\boldsymbol{M}_{ci} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -Var_{2,J}(y_i \mid x_i) \end{pmatrix} \tag{S1.15}$$

$$\frac{\partial \boldsymbol{d}_{qij}}{\partial \boldsymbol{\phi}_2} = \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))(1-2\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i}\boldsymbol{z}'_{2i}w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22} \tag{S1.16}$$

$$+ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \boldsymbol{z}'_{2i}\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\times \tag{S1.17}$$

$$\times \left\{[w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(q_{\tau_j}(x_i) - E_{2,J}(y_i \mid x_i))]\phi_{22} + w_{2ij}(\boldsymbol{\phi}_2, q_{\tau_j}(x_i))\right\} \tag{S1.18}$$

$$+ \boldsymbol{M}_{ci}\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22},$$

$$\check{\boldsymbol{d}}_{qij} = \boldsymbol{z}_{2i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})[\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\phi_{22}w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})] \tag{S1.19}$$

$$+ (\delta_{1i} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) \begin{pmatrix} 0 \\ 0 \\ w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(\phi_{22} + \frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))} - \phi_{22}E_{2,J}(Y \mid x_i; \phi_{22}, \boldsymbol{q}_{yi})) \end{pmatrix}$$

$$\frac{\partial \boldsymbol{d}_{qij}}{\partial q_{\tau_j}(x_i)} = \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))(1-2\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})^2 \phi_{22}^2(-1)$$

(S1.20)

$$+ \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\phi_{22}\times$$

$$\times \left\{ w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) \begin{pmatrix} 0 \\ 0 \\ w_{2ij}(1+\frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))}) \end{pmatrix} + \boldsymbol{z}_{2i} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}(1-w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) \right\}$$

(S1.21)

$$\frac{\partial \boldsymbol{d}_{qij}}{\partial q_{\tau_{j'}}(x_i)} = \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}^2 w_{2ij'}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(-1)$$

(S1.22)

$$+ \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\phi_{22}\times \quad \text{(S1.23)}$$

$$\times \left\{ w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) \begin{pmatrix} 0 \\ 0 \\ w_{2ij'}(1+\frac{q_{\tau_{j'}}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))}) \end{pmatrix} + \boldsymbol{z}_{2i} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) w_{2ij'}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(-1) \right\}$$

(S1.24)

$$\check{W}_{q\phi j\cdot,i} = \frac{\partial \boldsymbol{d}_{qij}}{\partial \boldsymbol{\phi}_2} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1-\pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))\boldsymbol{z}_{2i}(0, 0, w_{2ij}(1+\frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))})q_{\tau_j}(x_i)\phi_{22})$$

$$+ (\delta_{1i} - \pi_{12i})(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\boldsymbol{M}_{w1i}$$

$$\boldsymbol{M}_{w1i} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})[q_{\tau_j}(x_i) - E_{2,J}(y_i \mid x_i)] + \gamma_{1ij} \end{pmatrix} \tag{S1.25}$$

$$\begin{aligned} \gamma_{1ij} &= \frac{w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})[q_{\tau_j}(x_i) - E_{2,J}(y_i \mid x_i)]q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))} \\ &+ w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) \left[\frac{q_{\tau_j}(x_i)}{\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))} - \frac{\phi_{22} q_{\tau_j}(x_i) E_{2,J}(y_i \mid x_i)}{\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))} \right] \end{aligned} \tag{S1.26}$$

$$\begin{aligned} \check{\boldsymbol{W}}_{qqjj} &= \frac{\partial \boldsymbol{d}_{qij}}{\partial q_{\tau_j}(x_i)} + \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}\boldsymbol{z}_{2i} \\ &+ (\delta_{1i} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) \begin{pmatrix} 0 \\ 0 \\ w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}(1 - w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) + \gamma_{2ij} \end{pmatrix} \end{aligned} \tag{S1.27}$$

$$\gamma_{2ij} = w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}) \left(\frac{\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))\phi_{22} - q_{\tau_j}(x_i)\phi_{22}\exp(\phi_{22} q_{\tau_j}(x_i))\phi_{22}}{[\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))]^2} \right) \tag{S1.28}$$

$$+ \frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^J \exp(\phi_{22} q_{\tau_j}(x_i))} w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}(1 - w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))$$

$$\check{\boldsymbol{W}}_{qqjj'} = \frac{\partial \boldsymbol{d}_{qij}}{\partial q_{\tau_{j'}}(x_i)} \tag{S1.29}$$

$$+ \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi}))w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})\phi_{22}(0, 0, w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})(1 + \frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))})) \tag{S1.30}$$

$$+ (\delta_{1i} - \pi_{12i}(\boldsymbol{\phi}_2, \boldsymbol{q}_{yi})) \begin{pmatrix} 0 \\ 0 \\ \gamma_{3ij} \end{pmatrix}$$

$$\gamma_{3ij} = -w_{2ij}w_{2ij'} - w_{2ij}w_{2ij'}(1 + \frac{q_{\tau_j}(x_i)\phi_{22}}{\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))}) - w_{2ij}\frac{q_{\tau_j}(x_i)q_{\tau_{j'}}(x_i)\phi_{22}^2}{[\sum_{j=1}^{J}\exp(\phi_{22}q_{\tau_j}(x_i))]^2}$$

S1.4 Linear Approximation for $\hat{\boldsymbol{\phi}}_3$

A linear approximation for $\hat{\boldsymbol{\phi}}_3$ is

$$\hat{\boldsymbol{\phi}}_3 - \boldsymbol{\phi}_3 = \boldsymbol{I}_{\phi_3}^{-1}\left\{n^{-1}\sum_{i=1}^{n}\boldsymbol{U}_{\phi_3 i}\right\} + o_p(n^{-0.5}),$$

where $\boldsymbol{I}_{\phi_3} = \lim_{n\to\infty} \boldsymbol{I}_{n,\phi_3}$

$$\boldsymbol{I}_{n,\phi_3} = n^{-1} \sum_{i\in A_{13}} \pi_{13i}(\boldsymbol{\phi}_3, \boldsymbol{q}_{xi})(1 - \pi_{13i}(\boldsymbol{\phi}_3, \boldsymbol{q}_{xi}))\boldsymbol{z}_{3i}\boldsymbol{z}_{3i}',$$

$$\boldsymbol{U}_{\phi_3 i} = \delta_{3i}\boldsymbol{T}_{i\infty}(\boldsymbol{\phi}_3) + \delta_{1i}\left\{\boldsymbol{T}_{i\infty}(\boldsymbol{\phi}_3) + \sum_{j=1}^{J} \boldsymbol{F}_j \boldsymbol{m}_{ij}\right\}$$

$$\boldsymbol{T}_{i\infty}(\boldsymbol{\phi}_3) = (\delta_{1i} - \pi_{13\infty}(x_i))\boldsymbol{z}_{3i}, \boldsymbol{z}_{3i} = (-1, -E_{3,J}(x_i \mid y_i), -y_i)$$

$$\boldsymbol{F}_j = \lim_{n\to\infty} \frac{1}{n_1} \sum_{k=1}^{n} (\delta_{1k} + \delta_{3k})\boldsymbol{f}_{qkj}\boldsymbol{B}(x_k)', \boldsymbol{f}_{qij} = \pi_{13i}(\boldsymbol{\phi}_3, \boldsymbol{q}_{xi})(1 - \pi_{13i}(\boldsymbol{\phi}_3, \boldsymbol{q}_{xi}))w_{3ij}\phi_{31}\boldsymbol{z}_{3i}$$

$$\pi_{13i}(\phi_3, \boldsymbol{q}_{xi}) = \left\{1 + \exp\left[\phi_{30} + \log\left(J^{-1}\sum_{j=1}^{J} \exp\{\phi_{31}q_j(x_i)\}\right) + \phi_{32}y_i\right]\right\}^{-1}$$

$$\boldsymbol{m}_{kj} = \boldsymbol{H}_{n_1,x|y}(\tau_j)^{-1}\boldsymbol{B}(y_i)\psi_{\tau_j}(e_{x|y,i}(\tau_j)), \quad e_{x|y,i}(\tau_j) = x_i - \boldsymbol{B}(y_i)\hat{\boldsymbol{\beta}}_{x|y}(\tau_j)$$

$w_{3ij} = w_{3ij}(\boldsymbol{\phi}_3, \boldsymbol{q}_{xi})$, and $\boldsymbol{H}_{n_1,x|y}(\tau_j) = \lim_{n\to\infty} n_1^{-1} \sum_{i:\delta_i=1} f_{x|y,\delta=1}(q_{\tau_j}(y_i), y_i)\boldsymbol{B}(y_i)\boldsymbol{B}(y_i)' + \frac{\lambda_{n_1,x}}{n_1}\boldsymbol{D}_m'\boldsymbol{D}_m$. An estimator of the variance of $\hat{\boldsymbol{\phi}}_3$ is defined in a manner analogous to (S2.2) and is therefore deferred to supplementary material.

S1.5 Proof of Theorem 2

Proof of Theorem 2. As shorthand, let $\hat{q}_j(x_i) = \hat{q}_{\tau_j}(x_i)$ and $\hat{q}_j(y_i) = \hat{q}_{\tau_j}(y_i)$. We decompose the difference between the imputed estimator $\hat{\theta}$ and θ as

$$\sqrt{n}\{\hat{\theta} - \theta\} = \sqrt{n}(T_1 + T_{2y} + T_{2x} + T_{3y} + T_{3x} + T_{4y} + T_{4x}),$$

where

$$T_1 = \frac{1}{n}\sum_{i=1}^{n}(g(x_i, y_i) - Eg(X, Y))$$
$$T_{2y} = \frac{1}{n}\sum_{i=1}^{n}\delta_{2i}(E_2[g(x_i, Y) \mid x_i] - g(x_i, y_i)), T_{2x} = \frac{1}{n}\sum_{i=1}^{n}\delta_{3i}(E_3[g(X, y_i) \mid y_i] - g(x_i, y_i))$$

$$T_{3y} = \frac{1}{n}\sum_{i=1}^{n}\delta_{2i}\sum_{j=1}^{J}w_{2ij}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{xi})g(x_i, \hat{q}_j(x_i)) - \sum_{j=1}^{J}w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{xi})g(x_i, q_j(x_i)),$$
$$T_{3x} = \frac{1}{n}\sum_{i=1}^{n}\delta_{3i}\sum_{j=1}^{J}w_{3ij}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{yi})g(\hat{q}_j(y_i), y_i) - \sum_{j=1}^{J}w_{3ij}(\boldsymbol{\phi}_3, \boldsymbol{q}_{yi})g(q_j(y_i), y_i),$$
$$T_{4y} = \frac{1}{n}\sum_{i=1}^{n}\delta_{2i}(\sum_{j=1}^{J}w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{xi}))g(x_i, q_j(x_i)) - E_2[g(x_i, Y) \mid x_i]),$$

and $T_{4x} = n^{-1}\sum_{i=1}^{n}\delta_{3i}(\sum_{j=1}^{J}w_{3ij}(\boldsymbol{\phi}_3, \boldsymbol{q}_{yi}))g(q_j(y_i), y_i) - E_3[g(X, y_i) \mid y_i])$.

We further decompose T_{3y} as $T_{3y} = T_{3y1} + T_{3y2}$, where

$$T_{3y1} = \frac{1}{n}\sum_{i=1}^{n}\delta_{2i}\left[\sum_{j=1}^{J}w_{2ij}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{xi})g(x_i, \hat{\boldsymbol{q}}_j(x_i)) - \sum_{j=1}^{J}w_{2ij}(\boldsymbol{\phi}_2, \hat{\boldsymbol{q}}_{xi})g(x_i, \hat{q}_j(x_i))\right],$$

and $T_{3y2} = n^{-1}\sum_{i=1}^{n}\delta_{2i}\left[\sum_{j=1}^{J}w_{2ij}(\boldsymbol{\phi}_2, \hat{\boldsymbol{q}}_{xi})g(x_i, \hat{q}_j(x_i)) - \sum_{j=1}^{J}w_{2ij}(\boldsymbol{\phi}_2, \boldsymbol{q}_{xi})g(x_i, q_j(x_i))\right]$.

We expand T_{3y1} as

$$T_{3y1} = \left\{n^{-1}\sum_{i=1}^{n}\delta_{3i}Cov_2(g(x_i, Y), Y \mid x_i)\right\}(0, 0, 1)(\hat{\boldsymbol{\phi}}_2 - \boldsymbol{\phi}_2) + o_p(K_{n_1,y|x}^{0.5}n_1^{-0.5}),$$

and

$$
\begin{aligned}
T_{3y2} &= \frac{1}{n}\sum_{i=1}^{n}\delta_{2i}\left[\sum_{j=1}^{J}\tilde{c}_{ykj}(\hat{q}_{\tau_j}(x_i) - q_{\tau_j}(x_i))\right] \\
&= \frac{1}{n}\sum_{k=1}^{n}\delta_{2k}\sum_{j=1}^{J}\tilde{c}_{ykj}\left[\boldsymbol{B}(x_k)'\boldsymbol{H}_{n_1,y|x}(\tau_j)^{-1}\frac{1}{n_1}\sum_{i=1}^{n_1}\delta_{1i}\boldsymbol{B}(x_i)\psi_{\tau_j}(e_{y|x,i}(\tau_j)) + b^{a}_{\tau_j}(x_k) + b^{\lambda}_{\tau_j}(x_k)\right] \\
&= \frac{1}{n_1}\sum_{i=1}^{n}\delta_{1i}\sum_{j=1}^{J}\frac{1}{n}\sum_{k=1}^{n}\delta_{2k}\tilde{c}_{yij}\boldsymbol{B}(x_k)'\boldsymbol{H}_{n_1,y|x}(\tau_j)^{-1}\boldsymbol{B}(x_i)\psi_{\tau_j}(e_{y|x,i}(\tau_j)) \\
&+ \frac{1}{n}\sum_{k=1}^{n}\delta_{2k}\sum_{j=1}^{J}\tilde{c}_{ykj}\left[b^{a}_{\tau_j}(x_k) + b^{\lambda}_{\tau_j}(x_k)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\delta_{1i}\int_0^1\int_{M_{1x}}^{M_{2x}}\boldsymbol{C}_y(x,\tau)\boldsymbol{\ell}_i(\tau)dF(x \mid \delta = 2)d\tau + O_p(K_{n_1,y|x}^{-(p+2)}),
\end{aligned}
$$

where $\tilde{c}_{yij} = c_y(x_i, \tau_j)$. An analogous expansion holds for T_{3x}. The result then follows from Chen and Yu (2016), Theorem 5.4.3 of Fuller (2005), and arguments analogous to those used for the proof of Theorem 1. □

S2 Functional forms of estimators of variances

In the definitions of the variance estimators, we use $\hat{\boldsymbol{q}}_{yi}$ and $\hat{\boldsymbol{q}}_{xi}$ as snonyms for $\boldsymbol{y}_i^*$ and $\boldsymbol{x}_i^*$, respectively. We continue to use $\hat{q}_j(\nu)$ as shorthand for $\hat{q}_{\tau_j}(\nu)$, as in the proof of Theorem 2.

S2.1 Estimator of variance of $\hat{\boldsymbol{\phi}}_2$

An estimator of the variance of $\hat{\boldsymbol{\phi}}_2$ is

$$\hat{V}\{\hat{\boldsymbol{\phi}}_2\} = n^{-2}\hat{\boldsymbol{I}}_{n,\phi_2}^{-1}\left(\sum_{i=1}^{n}\hat{\boldsymbol{U}}_{\phi_2 i}\hat{\boldsymbol{U}}_{\phi_2 i}'\right)\hat{\boldsymbol{I}}_{n,\phi_2}^{-1}, \tag{S2.1}$$

where $\hat{\boldsymbol{I}}_{n,\phi_2} = n^{-1}\sum_{i\in A_{12}}\pi_{12i}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})(1-\pi_{12i}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi}))\hat{\boldsymbol{z}}_{2i}\hat{\boldsymbol{z}}_{2i}'$, $\hat{\boldsymbol{U}}_{\phi_2 i} = \delta_{2i}\hat{\boldsymbol{S}}_i(\hat{\boldsymbol{\phi}}_2)+\delta_{1i}[\hat{\boldsymbol{S}}_i(\hat{\boldsymbol{\phi}}_2)+\sum_{j=1}^{J}\hat{\boldsymbol{D}}_j\hat{\boldsymbol{\ell}}_{ij}]$, $\hat{\boldsymbol{S}}_i(\hat{\boldsymbol{\phi}}_2) = (\delta_{1i}-\pi_{12i}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi}))\hat{\boldsymbol{z}}_{2i}$, $\hat{\boldsymbol{z}}_{2i} = -(1, x_i, \hat{E}_{2,J}(y_i \mid x_i))$, $\hat{E}_{2,J} = \sum_{j=1}^{J} w_{2ij}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})\hat{q}_{\tau_j}(x_i)$,

$$\hat{\boldsymbol{D}}_j = \frac{1}{n_1}\sum_{k=1}^{n}(\delta_{1k}+\delta_{2k})\hat{\boldsymbol{d}}_{qkj}\boldsymbol{B}(x_k)'$$

$$\hat{\boldsymbol{d}}_{qij} = \pi_{12i}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})(1-\pi_{12i}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi}))w_{2ij}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})\hat{\phi}_{22}\hat{\boldsymbol{z}}_{2i}, \quad \hat{\boldsymbol{\ell}}_{ij} = \hat{\boldsymbol{H}}_{n_1,y|x}(\tau_j)\boldsymbol{B}(x_i)\psi_{\tau_j}(\hat{e}_{y|x,i}(\tau_j))$$

$$\hat{\boldsymbol{H}}_{n_1,y|x}(\tau_j) = \hat{\boldsymbol{\Phi}}_{y|x}(\tau)+\frac{\lambda_{n_1,y}}{n_1}\boldsymbol{D}_m'\boldsymbol{D}_m, \quad \hat{\boldsymbol{\Phi}}_{y|x}(\tau) = n_1^{-1}\sum_{i:\delta_i=1}\hat{f}_{y|x,\delta=1}(x_i,\hat{q}_\tau(x_i))\boldsymbol{B}(x_i)\boldsymbol{B}(x_i)'$$

$$\hat{f}_{y|x,\delta=1}(x,y) = \frac{(ab)^{-1}\sum_{i=1}^{n}\delta_{1i}K\left(\frac{y-y_i}{a}\right)K\left(\frac{x-x_i}{b}\right)}{a^{-1}\sum_{i=1}^{n}\delta_{1i}K\left(\frac{x-x_i}{a}\right)},$$

and $K(\cdot)$ is a Gaussian kernel with bandwidths a and b for x and y, respectively.

S2.2 Estimator of the Variance of $\hat{\boldsymbol{\phi}}_3$

An estimator of the variance of $\hat{\boldsymbol{\phi}}_3$ is

$$\hat{V}\{\hat{\boldsymbol{\phi}}_3\} = n^{-2}\hat{\boldsymbol{I}}_{n,\phi_3}^{-1}\left(\sum_{i=1}^{n}\hat{\boldsymbol{U}}_{\phi_3 i}\hat{\boldsymbol{U}}_{\phi_3 i}'\right)\hat{\boldsymbol{I}}_{n,\phi_3}^{-1}, \tag{S2.2}$$

where

$$\hat{\boldsymbol{I}}_{n,\boldsymbol{\phi}_3} = n^{-1} \sum_{i \in A_{13}} \pi_{13i}(\hat{\boldsymbol{\phi}}_3, \hat{q}_{xi})(1 - \pi_{13i}(\hat{\boldsymbol{\phi}}_3, \hat{q}_{xi}))\hat{\boldsymbol{z}}_{3i}\hat{\boldsymbol{z}}'_{3i}$$

$$\hat{\boldsymbol{U}}_{\boldsymbol{\phi}_3 i} = \delta_{3i}\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{\phi}}_3) + \delta_{1i}[\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{\phi}}_3) + \sum_{j=1}^{J} \hat{\boldsymbol{F}}_j \hat{\boldsymbol{m}}_{ij}],$$

$$\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{\phi}}_3) = (\delta_{1i} - \pi_{13i}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi}))\hat{\boldsymbol{z}}_{3i}$$

$$\hat{\boldsymbol{z}}_{3i} = -(1, y_i, \hat{E}_{2,J}(x_i \mid y_i))$$

$$\hat{E}_{2,J} = \sum_{j=1}^{J} w_{3ij}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi})\hat{q}_{\tau_j}(y_i)$$

$$\hat{\boldsymbol{F}}_j = \frac{1}{n_1} \sum_{k=1}^{n} (\delta_{1k} + \delta_{3k})\hat{\boldsymbol{f}}_{qkj}\boldsymbol{B}(y_k)'$$

$$\hat{\boldsymbol{f}}_{qij} = \pi_{13i}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi})(1 - \pi_{13i}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi}))w_{3ij}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi})\hat{\phi}_{31}\hat{\boldsymbol{z}}_{3i}$$

$$\hat{\boldsymbol{m}}_{ij} = \hat{\boldsymbol{H}}_{x|yn_1}(\tau_j)\boldsymbol{B}(y_i)\psi_{\tau_j}(\hat{e}_{x|y,i}(\tau_j))$$

$$\hat{\boldsymbol{H}}_{n_1 x|y}(\tau_j) = \hat{\boldsymbol{\Phi}}_{x|y}(\tau) + \frac{\lambda_{n_1,x}}{n_1}\boldsymbol{D}'_m\boldsymbol{D}_m,$$

$$\hat{\boldsymbol{\Phi}}_{x|y}(\tau) = n_1^{-1} \sum_{i:\delta_i=1} \hat{f}_{x|y,\delta=1}(y_i, \hat{q}_\tau(y_i))\boldsymbol{B}(y_i)\boldsymbol{B}(y_i)'$$

$$\hat{f}_{x|y,\delta=1}(x, y) = \frac{(ab)^{-1}\sum_{i=1}^{n} \delta_{1i} K\left(\frac{y-y_i}{a}\right) K\left(\frac{x-x_i}{b}\right)}{a^{-1}\sum_{i=1}^{n} \delta_{1i} K\left(\frac{x-x_i}{a}\right)},$$

where $K(\cdot)$ is a Gaussian kernel with bandwidths a and b for x and y, respectively.

S2.3 Definition of $\hat{r}_i$

An estimator of r_i is

$$\hat{r}_i = g(x_i, y_i) - \hat{\theta} + \delta_{2i}(\hat{E}_2[g(x_i, y_i) \mid x_i] - g(x_i, y_i)) + \delta_{3i}(\hat{E}_3[g(x_i, y_i) \mid y_i] - g(x_i, y_i)) \tag{S2.3}$$

$$+ (\delta_{1i} + \delta_{2i}) \left\{ n^{-1} \sum_{k=1}^{n} \delta_{2k} \hat{Cov}_2(g(x_k, Y), Y \mid x_k) \right\} \boldsymbol{e}_2' \hat{\boldsymbol{I}}_{n,\phi_2}^{-1} \hat{\boldsymbol{U}}_{\phi_2,i} \tag{S2.4}$$

$$+ (\delta_{1i} + \delta_{3i}) \left\{ n^{-1} \sum_{k=1}^{n} \delta_{3k} \hat{Cov}_3(g(X, y_k), X \mid y_k) \right\} \boldsymbol{e}_3' \hat{\boldsymbol{I}}_{n,\phi_3}^{-1} \hat{\boldsymbol{U}}_{\phi_3,i}$$

$$+ \delta_{1i}(\sum_{j=1}^{J} \hat{\boldsymbol{C}}_{yj}' \hat{\boldsymbol{\ell}}_{ij} + \hat{\boldsymbol{C}}_{xj}' \hat{\boldsymbol{m}}_{ij}) \tag{S2.5}$$

$\boldsymbol{e}_3 = (0, 0, 1)'$, $\hat{\boldsymbol{C}}_{yj} = n_1^{-1} \sum_{i=1}^{n} \delta_{2i} \hat{\boldsymbol{C}}_{yij}$, $\hat{\boldsymbol{C}}_{xj} = n_1^{-1} \sum_{i=1}^{n} \delta_{3i} \hat{\boldsymbol{C}}_{xij}$, $\hat{\boldsymbol{C}}_{yij}' = \hat{\tilde{c}}_{yij} \boldsymbol{B}(x_i)'$, $\hat{\boldsymbol{C}}_{xij}' = \hat{\tilde{c}}_{xij} \boldsymbol{B}(y_i)'$,

$$\hat{\tilde{c}}_{yij} = \frac{\hat{c}_{yij}}{\sum_{j=1}^{J} \exp(\hat{\phi}_{22} \hat{q}_j(x_i))} - \hat{E}_2[g(x_i, Y) \mid x_i] w_{2ij}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{yi}) \hat{\phi}_{22}$$

$$\hat{\tilde{c}}_{xij} = \frac{\hat{c}_{xij}}{\sum_{j=1}^{J} \exp(\hat{\phi}_{31} \hat{q}_j(y_i))} - \hat{E}_3[g(X, y_i) \mid y_i] w_{3ij}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi}) \hat{\phi}_{31}$$

$\hat{c}_{yij} = \exp(\hat{\phi}_{22} \hat{q}_j(x_i)) g_y'(x_{ij}, \hat{q}_j(x_i)) + g(x_i, \hat{q}_j(x_i)) \exp(\hat{\phi}_{22} \hat{q}_j(x_i)) \hat{\phi}_{22}$, and $\hat{c}_{xij} = \exp(\hat{\phi}_{31} \hat{q}_j(y_i)) g_x'(\hat{q}_j(y_i), y_i) + g(\hat{q}_j(y_i), y_i) \exp(\hat{\phi}_{31} \hat{q}_j(y_i)) \hat{\phi}_{31}$. The estimated moments are $\hat{E}_2[g(x_i, y_i) \mid x_i] = J^{-1} \sum_{j=1}^{J} w_{2ij}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{yi}) g(x_i, \hat{q}_j(x_i))$, $\hat{E}_3[g(x_i, y_i) \mid$

$y_i] = J^{-1}\sum_{j=1}^{J} w_{3ij}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xi})g(\hat{q}_j(y_i), y_i)$,

$$\hat{Cov}_2(g(x_k, Y), Y \mid x_k) = J^{-1}\sum_{j=1}^{J} w_{2kj}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{yk})g(x_k, \hat{q}_j(x_k))\hat{q}_j(x_k) - J^{-1}\sum_{j=1}^{J} w_{2kj}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{yk})g(x_k, \hat{q}_j(x_k))J^{-1}\sum_{j=1}^{J} w_{2kj}(\hat{\boldsymbol{\phi}}_2, \hat{\boldsymbol{q}}_{yk})\hat{q}_j(x_k),$$

and

$$\hat{Cov}_3(g(X, y_k), X \mid y_k) = J^{-1}\sum_{j=1}^{J} w_{3kj}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xk})g(\hat{q}_j(y_k), y_k)\hat{q}_j(y_k) - J^{-1}\sum_{j=1}^{J} w_{3kj}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xk})g(\hat{q}_j(y_k), y_k)J^{-1}\sum_{j=1}^{J} w_{3kj}(\hat{\boldsymbol{\phi}}_3, \hat{\boldsymbol{q}}_{xk})\hat{q}_j(y_k)$$

S2.4 Composite Estimators

One may be interested in a parameter of the form $\theta = h(\theta_1, \ldots, \theta_K)$, where each θ_k is of the form $Eg_k(X, Y)$, for some function $g_k(X, Y)$. For example, if the parameter of interest is $Cov(X, Y)$, then $\theta = \theta_1 - \theta_2\theta_3$, where $\theta_1 = EXY$, $\theta_2 = EX$, and $\theta_3 = EY$. The estimator of θ is of the form $\hat{\theta} = h(\hat{\theta}_1, \ldots, \hat{\theta}_K)$. An estimator of the variance of $\hat{\theta}$ is given by $\hat{V}\{\hat{\theta}\} = n^{-2}\hat{\boldsymbol{d}}_h' \sum_{i=1}^{n}(\hat{r}_{i1}, \ldots, \hat{r}_{iK})'(\hat{r}_{i1}, \ldots, \hat{r}_{iK})\hat{\boldsymbol{d}}_h$, where

$$\hat{\boldsymbol{d}}_h = \left(\frac{\partial h}{\partial \theta_1}, \ldots, \frac{\partial h}{\partial \theta_K}\right) |_{(\hat{\theta}_1, \ldots, \hat{\theta}_K)},$$

and $\hat{r}_{ik}$ is the estimator of r_i of (3.4) of the main manuscript, appropriate for estimating the variance of $\hat{\theta}_k$.

S3 Propensity Score Adjusted Imputed Estimator: Further Detail

S3.1 Variance Estimator for Propensity Score Adjusted Imputed Estimator

The estimator of the variance of the propensity-score adjusted imputed estimator is

$$\hat{V}_{PSA-IMP}(\hat{\theta}_{PSA-IMP}) = \frac{1}{n^2}\sum_{i=1}^{n}(\hat{r}_{i,p} - \bar{r}_p)^2,$$

where $\bar{r}_p = n^{-1}\sum_{i=1}^{n}\hat{r}_{i,p}$,

$$\hat{r}_{i,p} = \hat{d}_{4i} + \hat{p}_{4i}^{-1}(g(x_i, y_i) - \hat{\theta}) + \hat{p}_{4i}^{-1}\delta_{2i}(\hat{E}_2[g(x_i, y_i) \mid x_i] - g(x_i, y_i)) + \hat{p}_{4i}^{-1}\delta_{3i}(\hat{E}_3[g(x_i, y_i) \mid y_i] - g(x_i, y_i)) \tag{S3.6}$$

$$+ (\delta_{1i} + \delta_{2i})\left\{n^{-1}\sum_{k=1}^{n}\delta_{2k}\hat{p}_{4k}^{-1}\hat{Cov}_2(g(x_k, Y), Y \mid x_k)\right\}\boldsymbol{e}_2'\hat{\boldsymbol{I}}_{n,\phi_2}^{-1}\hat{\boldsymbol{U}}_{\phi_2,i} \tag{S3.7}$$

$$+ (\delta_{1i} + \delta_{3i})\left\{n^{-1}\sum_{k=1}^{n}\delta_{3k}\hat{p}_{4k}^{-1}\hat{Cov}_3(g(X, y_k), X \mid y_k)\right\}\boldsymbol{e}_3'\hat{\boldsymbol{I}}_{n,\phi_3}^{-1}\hat{\boldsymbol{U}}_{\phi_3,i}$$

$$+ \delta_{1i}(\sum_{j=1}^{J}\hat{\boldsymbol{C}}_{yj,p}'\hat{\boldsymbol{\ell}}_{ij} + \hat{\boldsymbol{C}}_{xj,p}'\hat{\boldsymbol{m}}_{ij}) \tag{S3.8}$$

$$\hat{d}_{4i} = \bar{\boldsymbol{V}}_{4,g}\boldsymbol{I}_{n,4}^{-1}(1,\boldsymbol{v}_i')'(1-\delta_{4i}-\hat{p}_{4i}),$$

$$\bar{\boldsymbol{V}}_{4,g} = \frac{1}{n}\left\{\sum_{i=1}^{n}\frac{g(x_i,y_i)}{\hat{p}_{4i}^2} + \sum_{j=1}^{J} w_{2ij}(\hat{\boldsymbol{\phi}}_2)\frac{g(x_i,y_{ij}^*)}{\hat{p}_{4i}^2} + \sum_{j=1}^{J} w_{3ij}(\hat{\boldsymbol{\phi}}_3)\frac{g(x_{ij}^*,y_i)}{\hat{p}_{4i}^2}\right\}(1,\boldsymbol{v}_i')'$$

$$\boldsymbol{I}_{n,4} = \sum_{i=1}^{n}(1,\boldsymbol{v}_i')'(1,\boldsymbol{v}_i')\hat{p}_{4i}(1-\hat{p}_{4i})$$

$\boldsymbol{e}_3 = (0,0,1)'$, $\hat{\boldsymbol{C}}_{yj,p} = n_1^{-1}\sum_{i=1}^{n}\delta_{2i}\hat{p}_{4i}^{-1}\hat{\boldsymbol{C}}_{yij}$, $\hat{\boldsymbol{C}}_{xj} = n_1^{-1}\sum_{i=1}^{n}\delta_{3i}\hat{p}_{4i}^{-1}\hat{\boldsymbol{C}}_{xij}$,

$\hat{\boldsymbol{C}}_{yij}' = \hat{\hat{c}}_{yij}\boldsymbol{B}(x_i)'$, $\hat{\boldsymbol{C}}_{xij}' = \hat{\hat{c}}_{xij}\boldsymbol{B}(y_i)'$,

$$\hat{\hat{c}}_{yij} = \frac{\hat{c}_{yij}}{\sum_{j=1}^{J}\exp(\hat{\phi}_{22}\hat{q}_j(x_i))} - \hat{E}_2[g(x_i,Y)\mid x_i]w_{2ij}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})\hat{\phi}_{22}$$

$$\hat{\hat{c}}_{xij} = \frac{\hat{c}_{xij}}{\sum_{j=1}^{J}\exp(\hat{\phi}_{31}\hat{q}_j(y_i))} - \hat{E}_3[g(X,y_i)\mid y_i]w_{3ij}(\hat{\boldsymbol{\phi}}_3,\hat{\boldsymbol{q}}_{xi})\hat{\phi}_{31}$$

$\hat{c}_{yij} = \exp(\hat{\phi}_{22}\hat{q}_j(x_i))g_y'(x_i,\hat{q}_j(x_i)) + g(x_i,\hat{q}_j(x_i))\exp(\hat{\phi}_{22}\hat{q}_j(x_i))\hat{\phi}_{22}$, and $\hat{c}_{xij} = \exp(\hat{\phi}_{31}\hat{q}_j(y_i))g_x'(\hat{q}_j(y_i),y_i) + g(\hat{q}_j(y_i),y_i)\exp(\hat{\phi}_{31}\hat{q}_j(y_i))\hat{\phi}_{31}$. The estimated moments are $\hat{E}_2[g(x_i,y_i)\mid x_i] = J^{-1}\sum_{j=1}^{J}w_{2ij}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yi})g(x_i,\hat{q}_j(x_i))$, $\hat{E}_3[g(x_i,y_i)\mid y_i] = J^{-1}\sum_{j=1}^{J}w_{3ij}(\hat{\boldsymbol{\phi}}_3,\hat{\boldsymbol{q}}_{xi})g(\hat{q}_j(y_i),y_i)$,

$$\hat{Cov}_2(g(x_k,Y),Y\mid x_k) = J^{-1}\sum_{j=1}^{J}w_{2kj}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yk})g(x_k,\hat{q}_j(x_k))\hat{q}_j(x_k) - J^{-1}\sum_{j=1}^{J}w_{2kj}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yk})g(x_k,\hat{q}_j(x_k))J^{-1}\sum_{j=1}^{J}w_{2kj}(\hat{\boldsymbol{\phi}}_2,\hat{\boldsymbol{q}}_{yk})\hat{q}_j(x_k),$$

and

$$\hat{Cov}_3(g(X,y_k),X\mid x_k) = J^{-1}\sum_{j=1}^{J}w_{3kj}(\hat{\boldsymbol{\phi}}_3,\hat{\boldsymbol{q}}_{xk})g(\hat{q}_j(y_k),y_k)\hat{q}_j(y_k) - J^{-1}\sum_{j=1}^{J}w_{3kj}(\hat{\boldsymbol{\phi}}_3,\hat{\boldsymbol{q}}_{xk})g(\hat{q}_j(y_k),y_k)J^{-1}\sum_{j=1}^{J}w_{3kj}(\hat{\boldsymbol{\phi}}_3,\hat{\boldsymbol{q}}_{xk})\hat{q}_j(y_k)$$

S3.2 Simulation for Propensity Score Adjustment

We verify that the propensity-score adjusted imputed estimator and corresponding variance estimator are unbiased through simulation. In the model for the simulation,

$$P(\delta_{4i}=1)=\frac{\exp(\phi_{40}+\phi_{41}v_i)}{1+\exp(\phi_{40}+\phi_{41}v_i)}:=p_{4i}$$

where $v_i \overset{ii}{\sim} Unif(-1,1)$ for $i=1,\ldots,n$, $\phi_{40}=-1.2$ and $\phi_{41}=0.2$. We generate $\delta_{4i} \overset{ind}{\sim} Bernoulli(p_{4i})$ for $i=1,\ldots,n$. We let $x_i=0.5+0.5v_i+e_i$, where $e_i \overset{iid}{\sim} N(0,0.3^2)$ for $i=1,\ldots,n$. We generate y_i as

$$y_i=\exp(2(x_i-0.5)/1.5)+\epsilon_i,$$

where $\epsilon_i \overset{iid}{\sim} N(0,0.2^2)$. We set $n=2000$, $J=50$, $\lambda=2$, and use 35 evenly-spaced knots.

Table 1 summarizes the MC properties of the estimator and variance estimator for a MC sample size of 100. The population parameter θ is the average across the 100 MC samples. From MC mean of the PS estimator in the row labeled $E_{MC}(\hat{\theta}_{PS})$, we conclude that the PS estimator is approximately unbiased for θ. The rows $V_{MC}(\hat{\theta}_{PS})$ and $E_{MC}(\hat{V}_{PS})$ contain the MC variance of $\hat{\theta}_{PS}$ an the MC mean of the variance estimator, respectively. The variance estimator slightly over-estimates the variance of the estimator. The t-statistics in the final row of the table indicate that the over-estimation of the variance is not significant relative to the MC error. Further, the coverage rates in the row labeled C.R.% are close to the nominal 95% level.

	EY	EX	EY^2	EX^2	EXY
θ	1.1655	0.4994	1.8667	0.4236	0.8493
$E_{MC}(\hat{\theta}_{PS})$	1.1655	0.5003	1.8651	0.4235	0.8489
$V_{MC}(\hat{\theta}_{PS})$	0.0399	0.0126	0.4936	0.0166	0.0842
$T(\hat{V})$	0.0372	0.0127	0.5137	0.0161	0.0830
C.R. %	0.9500	0.9400	0.9500	0.9400	0.9300

Table 1: MC properties of propensity-score adjusted imputed estimator and corresponding variance estimator.

S3.3 Covariate Definitions for Propensity Score Model in the Data Analysis

The variable Age is the age of the head of the household, with 5 possible values: 1= under 30 years, 2=30-39, 3=40-49, 4=50-59, 5=60 and older. Income represents the household income with 5 possible values: 1=under $20,000; 2=$20,000 to $39,999; 3=$40,000 to $59,999; 4=$60,000 to $99,999; 5=$100,000 and over. Educational status is similarly summarized through a numeric variable, where 1= high school or less, 2=attended college, 3=college graduate, 4=advanced degree. Finally, the household size has possible values of 1 person, 2 people, and 3 for more than 2 people.

S4 Identification Condition for Multivariate Covariates

Consider a $p-$dimensional vector of covariates $\boldsymbol{x} = (x_1, \ldots, x_p)$. A multivariate version of the B-spline is $\boldsymbol{B}(\boldsymbol{x}) = (\boldsymbol{B}_1(x_1), \ldots, B_p(x_p))'$, where $\boldsymbol{B}_k(x_k)$ is the B-spline for univariate variable x_k. Assume all elements of $\boldsymbol{x}$ are simultaneously observed or are all missing so that Δ retains its interpretation. Assume $P(\Delta = k) \propto \exp(\phi_{k0} + \boldsymbol{\phi}'_{k1}\boldsymbol{x} + \phi_{k2}y)$ such that $\sum_{k=1}^{3} P(\Delta = k) = 1$. Define $h_x(\boldsymbol{\phi}_{31}, y) = -\log(E[\exp(-\boldsymbol{\phi}'_{31}\boldsymbol{X}) \mid y, \delta = 1])$. To identify the parameters, one must ensure that the matrix $\boldsymbol{V}$ is full rank, where $\boldsymbol{V} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)'$, and $\boldsymbol{v}_i = (y_i, \partial h_x(\boldsymbol{\phi}_{31}, y)/\partial \boldsymbol{\phi}_{31})$.

References

Berg, E., Kim, J. K., & Skinner, C. (2016). Imputation under informative sampling. *Journal of Survey Statistics and Methodology*, **4**, 436-462.

Chen, S. and Yu, C.L. (2016). Parameter estimation through semiparametric quantile regression imputation. *Electron. J. Statist.*, **10**, 3621–3647.

Fuller, W.A. (1996). *Introduction to Statistical Time Series*, John Wiley and Sons, New York.

REFERENCES

YOSHIDA, T. (2013) Asymptotics for Penalized Spline Estimators in Quantile Regression. *Communications in Statistics - Theory and Methods*.