IMPROVED MODEL-ASSISTED ESTIMATION VIA PROBABILITY THRESHOLDING

Xianpeng Zong, Geoffrey K. F. Tso and Guohua Zou

Beijing University of Technology, City University of Hong Kong, and Capital Normal University

Supplementary Material

This supplementary material contains the proofs of theoretical results and additional simulation results.

S1 Proofs of theoretical results

Denote C as a generic positive constant which may be different in different places.

S1.1 Proof of Theorem 1

Proof. From Theorem 1 of Zong et al. (2019), C_r inequality and Conditions 1 and 2, we observe that

$$\mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{IHT}}) = \frac{1}{N^2}\left\{\left(\sum_{k\in U_2}\frac{\pi_k-\pi_k^*}{\pi_k^*}y_k\right)^2+\sum_{k\in U}\frac{\pi_k(1-\pi_k)}{\pi_k^{*2}}y_k^2+\sum_{k\neq l\in U}\frac{\pi_{kl}-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right\}$$

$$\leq \frac{K}{N^2}\sum_{k\in U_2} y_k^2 + \frac{1}{N^2\pi_{(K)}^2}\sum_{k\in U}\pi_k(1-\pi_k)y_k^2$$
$$+ \max_{k\neq l\in U}|\pi_{kl}-\pi_k\pi_l|\cdot\frac{1}{N^2\pi_{(K)}^2}\sum_{k\neq l\in U}|y_k y_l|$$
$$\leq \frac{CK^2}{N^2} + \frac{C}{N\pi_{(K)}^2} + \frac{C}{n\pi_{(K)}^2}. \tag{S1.1}$$

Note that $\pi_{(K)} \geq \lambda_N$. Hence, to verify Theorem 1, it suffices to show that the correction ratio $K/N = O\left(n^{-1/2}\lambda_N^{-1}\right)$.

By Definition 1, we have $\lambda_N \leq \pi_{(K)} \leq (K+1)^{-1}$, and so $K \leq \lambda_N^{-1}$. Thus, $\sqrt{n}\lambda_N K/N \leq \sqrt{n}/N \leq C$. This completes the proof of Theorem 1.

S1.2 Proof of Theorem 2

Proof. Following Theorem 1 of Zong et al. (2019), we have

$$\mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{IHT}}) = \frac{1}{N^2}\left\{\left(\sum_{k\in U_2}\frac{\pi_k-\pi_k^*}{\pi_k^*}y_k\right)^2 + \sum_{k\in U}\frac{\pi_k(1-\pi_k)}{\pi_k^{*2}}y_k^2\right\}$$
$$+\frac{1}{N^2}\sum_{k\neq l\in U}\frac{\pi_{kl}-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_k y_l$$
$$\triangleq F_1 + F_2.$$

From Equation (2.2), the MSE_p of HT estimator is

$$\mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{HT}}) = \frac{1}{N^2}\sum_{k\in U}\frac{1-\pi_k}{\pi_k}y_k^2 + \frac{1}{N^2}\sum_{k\neq l\in U}\frac{\pi_{kl}-\pi_k\pi_l}{\pi_k\pi_l}y_k y_l$$
$$\triangleq F_3 + F_4.$$

Note that

$$
\begin{aligned}
F_3 - F_1 &= \frac{1}{N^2}\sum_{k\in U}\frac{1-\pi_k}{\pi_k}y_k^2 - \frac{1}{N^2}\left\{\left(\sum_{k\in U_2}\frac{\pi_k-\pi_k^*}{\pi_k^*}y_k\right)^2 + \sum_{k\in U}\frac{\pi_k(1-\pi_k)}{\pi_k^{*2}}y_k^2\right\}\\
&= \frac{1}{N^2}\left\{\sum_{k\in U_2}\frac{(\pi_k^{*2}-\pi_k^2)(1-\pi_k)}{\pi_k^{*2}\pi_k}y_k^2 - \left(\sum_{k\in U_2}\frac{\pi_k-\pi_k^*}{\pi_k^*}y_k\right)^2\right\}\\
&\geq \frac{1}{N^2}\left\{\sum_{k\in U_2}\frac{(\pi_k^{*2}-\pi_k^2)(1-\pi_k)}{\pi_k^{*2}\pi_k}y_k^2 - K\sum_{k\in U_2}\frac{(\pi_k-\pi_k^*)^2}{\pi_k^{*2}}y_k^2\right\}\\
&= \frac{1}{N^2}\sum_{k\in U_2}\frac{(\pi_k^*-\pi_k)[(1-\pi_k-K\pi_k)\pi_k^*+(\pi_k-\pi_k^2+K\pi_k^2)]}{\pi_k^{*2}\pi_k}y_k^2.
\end{aligned}
$$

From Definition 1, we have $\pi_k \leq \pi_k^* \leq (K+1)^{-1}$ for each $k \in U_2$. Thus $F_3 - F_1 \geq 0$.

For the terms F_2 and F_4, by Condition 2, we have

$$
\begin{aligned}
|F_4 - F_2| &= \frac{1}{N^2}\left|\sum_{k\neq l\in U}\frac{\pi_{kl}-\pi_k\pi_l}{\pi_k\pi_l}y_ky_l - \sum_{k\neq l\in U}\frac{\pi_{kl}-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right|\\
&\leq \max_{k\neq l\in U}|\pi_{kl}-\pi_k\pi_l|\cdot\frac{1}{N^2}\sum_{k\neq l\in U}\left|\frac{\pi_k^*\pi_l^*-\pi_k\pi_l}{\pi_k\pi_l\pi_k^*\pi_l^*}y_ky_l\right|\\
&\leq \frac{C}{nN^2\lambda_N^2}\cdot\sum_{k\neq l\in U}\left|\frac{\pi_k^*\pi_l^*-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right|. \qquad \text{(S1.2)}
\end{aligned}
$$

Further, from Condition 1 and the definition of π_k^*, it follows that

$$
\begin{aligned}
\sum_{k\neq l\in U}\left|\frac{\pi_k^*\pi_l^*-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right| &= K^2\cdot\frac{1}{K^2}\sum_{k\neq l\in U_2}\left|\frac{\pi_k^*\pi_l^*-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right|\\
&\quad + NK\cdot\frac{2}{NK}\sum_{k\in U_1, l\in U_2}\left|\frac{\pi_k^*\pi_l^*-\pi_k\pi_l}{\pi_k^*\pi_l^*}y_ky_l\right|\\
&\leq CNK. \qquad \text{(S1.3)}
\end{aligned}
$$

Combining Equations (S1.2) and (S1.3), we obtain $|F_4 - F_2| \leq CK(Nn\lambda_N^2)^{-1}$.

As in the proof of Theorem 1, we have $K \le \lambda_N^{-1}$. So $|F_4 - F_2| \le C(nN\lambda_N^3)^{-1}$. Thus, by the condition $(N\lambda_N^3)^{-1} = o(1)$, we see that

$$\mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{IHT}}) \le \mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{HT}}) + o\left(n^{-1}\right).$$

Especially, for Poisson sampling, we have $\pi_{kl} = \pi_k \pi_l$, and in this case, $F_2 = F_4 = 0$. Consequently,

$$\mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{IHT}}) \le \mathrm{MSE}_p(\hat{\hat{t}}_{\mathrm{HT}}),$$

where the strict inequality holds if and only if there exist $k \ne l \in U_2$ such that $(\pi_k - \pi_{(K)})y_k \ne (\pi_l - \pi_{(K)})y_l$. This completes the proof of Theorem 2.

S1.3 Lemmas for the proofs of Theorems 3 - 7

Lemma 1. *Let g_{ijkl} be an arbitrary non-random variable. If Conditions 2 and 4 hold, then*

$$\frac{1}{N^4}\sum_{i,j,k,l\in U}\left|g_{ijkl}\cdot E_p\left(\Phi_i\Phi_j\Phi_k\Phi_l\right)\right| \le \frac{C}{n^2\lambda_N^4}\cdot\max_{i,j,k,l\in U}|g_{ijkl}|$$

and

$$\frac{1}{N^4}\sum_{i,j,k,l\in U}\left|g_{ijkl}\cdot E_p\left(\Phi_i^*\Phi_j^*\Phi_k^*\Phi_l^*\right)\right| \le \frac{C}{n^2\pi_{(K)}^4}\cdot\max_{i,j,k,l\in U}|g_{ijkl}|,$$

where $\Phi_i = (I_i - \pi_i)/\pi_i$ and $\Phi_i^* = (I_i - \pi_i^*)/\pi_i^*$.

Proof. From Condition 2 and the first assumption of Condition 4, we have

$$\max_{(i,j,k)\in D_{3,N}}|\mathrm{E}_p[(I_i-\pi_i)(I_j-\pi_j)(I_k-\pi_k)]| = O(n^{-1})$$

and

$$\max_{(i,j,k)\in D_{3,N}} \left|\mathrm{E}_p[(I_i-\pi_i)^2(I_j-\pi_j)(I_k-\pi_k)]\right| = O(n^{-1}).$$

Then we observe that

$$\begin{aligned}
&\frac{1}{N^4}\sum_{i,j,k,l\in U}\left|g_{ijkl}\cdot \mathrm{E}_p\left(\Phi_i\Phi_j\Phi_k\Phi_l\right)\right| \\
&\le \frac{1}{N^4\lambda_N^4}\cdot \max_{i,j,k,l\in U}|g_{ijkl}|\cdot \sum_{i,j,k,l\in U}|\mathrm{E}_p[(I_i-\pi_i)(I_j-\pi_j)(I_k-\pi_k)(I_l-\pi_l)]| \\
&= \frac{1}{N^4\lambda_N^4}\cdot \max_{i,j,k,l\in U}|g_{ijkl}|\cdot \left\{\sum_{i,j,k,l\in D_{4,N}}|\mathrm{E}_p[(I_i-\pi_i)(I_j-\pi_j)(I_k-\pi_k)(I_l-\pi_l)]|\right. \\
&\quad + \sum_{i,j,k\in D_{3,N}}\left|\mathrm{E}_p[(I_i-\pi_i)^2(I_j-\pi_j)(I_k-\pi_k)]\right| + \sum_{i\in U}\left|\mathrm{E}_p(I_i-\pi_i)^4\right| \\
&\quad \left. + \sum_{i\neq j\in U}\left|\mathrm{E}_p[(I_i-\pi_i)^2(I_j-\pi_j)^2+(I_i-\pi_i)(I_j-\pi_j)^3]\right|\right\} \\
&\le \frac{C}{n^2\lambda_N^4}\cdot \max_{i,j,k,l\in U}|g_{ijkl}|. \qquad \text{(S1.4)}
\end{aligned}$$

Note that

$$\Phi_i^* = \frac{I_i-\pi_i}{\pi_i^*}+\frac{\pi_i-\pi_i^*}{\pi_i^*} \triangleq \tilde{\Phi}_i+\Delta_i.$$

Then

$$\begin{aligned}
&\frac{1}{N^4}\sum_{i,j,k,l\in U}\left|g_{ijkl}\cdot \mathrm{E}_p\left(\Phi_i^*\Phi_j^*\Phi_k^*\Phi_l^*\right)\right| \\
&\le \frac{1}{N^4}\cdot \max_{i,j,k,l\in U}|g_{ijkl}|\cdot \sum_{i,j,k,l\in U}\left|\mathrm{E}_p\left[\left(\tilde{\Phi}_i+\Delta_i\right)\left(\tilde{\Phi}_j+\Delta_j\right)\left(\tilde{\Phi}_k+\Delta_k\right)\right.\right. \\
&\quad \left.\left.\left(\tilde{\Phi}_l+\Delta_l\right)\right]\right|
\end{aligned}$$

$$\leq \frac{C}{N^4} \cdot \max_{i,j,k,l \in U} |g_{ijkl}| \cdot \sum_{i,j,k,l \in U} \Big\{ \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k \tilde{\Phi}_l) \Big| + \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k \Delta_l) \Big|$$

$$+ \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \Delta_k \Delta_l) \Big| + \Big| \mathrm{E}_p(\tilde{\Phi}_i \Delta_j \Delta_k \Delta_l) \Big| + |\mathrm{E}_p(\Delta_i \Delta_j \Delta_k \Delta_l)| \Big\}$$

$$= \frac{C}{N^4} \cdot \max_{i,j,k,l \in U} |g_{ijkl}| \cdot \left\{ \left(\sum_{l \in U} |\Delta_l| \right)^4 + \left(\sum_{l \in U} |\Delta_l| \right) \left(\sum_{i,j,k \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k) \Big| \right) \right.$$

$$\left. + \left(\sum_{l \in U} |\Delta_l| \right)^2 \left(\sum_{i,j \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j) \Big| \right) + \sum_{i,j,k,l \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k \tilde{\Phi}_l) \Big| \right\}. \quad \text{(S1.5)}$$

Similar to (S1.4), it is clear that

$$\frac{1}{N^4} \sum_{i,j,k,l \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k \tilde{\Phi}_l) \Big|$$

$$\leq \frac{1}{N^4 \pi_{(K)}^4} \sum_{i,j,k,l \in U} |\mathrm{E}_p[(I_i - \pi_i)(I_j - \pi_j)(I_k - \pi_k)(I_l - \pi_l)]|$$

$$= O\left(\frac{1}{n^2 \pi_{(K)}^4} \right). \quad \text{(S1.6)}$$

From Conditions 2 and 4, we obtain that

$$\frac{1}{N^2} \sum_{i,j \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j) \Big| \leq \frac{1}{N^2 \pi_{(K)}^2} \left\{ \sum_{i \neq j \in U} |\mathrm{E}_p[(I_i - \pi_i)(I_j - \pi_j)]| \right.$$

$$\left. + \sum_{i \in U} \left| \mathrm{E}_p(I_i - \pi_i)^2 \right| \right\}$$

$$= O\left(\frac{1}{n \pi_{(K)}^2} \right) \quad \text{(S1.7)}$$

and

$$\frac{1}{N^3} \sum_{i,j,k \in U} \Big| \mathrm{E}_p(\tilde{\Phi}_i \tilde{\Phi}_j \tilde{\Phi}_k) \Big|$$

$$\leq \frac{C}{N^3\pi_{(K)}^3}\left\{\sum_{i,j,k\in D_{3,N}} |\mathrm{E}_p[(I_i-\pi_i)(I_j-\pi_j)(I_k-\pi_k)]|\right.$$
$$\left. + \sum_{i\neq j\in U}\left|\mathrm{E}_p[(I_i-\pi_i)^2(I_j-\pi_j)]\right| + \sum_{i\in U}\left|\mathrm{E}_p(I_i-\pi_i)^3\right|\right\}$$
$$= O\left(\frac{1}{n\pi_{(K)}^3}\right). \tag{S1.8}$$

By Definition 1, we have $K \leq \pi_{(K)}^{-1}$, and so

$$\frac{1}{N}\sum_{l\in U}|\Delta_l| = \frac{K}{N}\left(\frac{1}{K}\sum_{l\in U_2}\left|\frac{\pi_l-\pi_l^*}{\pi_l^*}\right|\right) = O\left(\frac{1}{N\pi_{(K)}}\right). \tag{S1.9}$$

Now combining Equations (S1.5) - (S1.9), we see that

$$\frac{1}{N^4}\sum_{i,j,k,l\in U}\left|g_{ijkl}\cdot \mathrm{E}_p\left(\Phi_i^*\Phi_j^*\Phi_k^*\Phi_l^*\right)\right| \leq \frac{C}{n^2\pi_{(K)}^4}\cdot \max_{i,j,k,l\in U}|g_{ijkl}|\,.$$

This completes the proof of Lemma 1.

Lemma 2. *If Conditions 1, 2 and 4 hold, then*

$$E_p(\hat{\bar{t}}_{HT}-\bar{t}_y)^4 = O\left(\frac{1}{n^2\lambda_N^4}\right) \quad \textit{and} \quad E_p(\hat{\bar{t}}_{IHT}-\bar{t}_y)^4 = O\left(\frac{1}{n^2\pi_{(K)}^4}\right).$$

Proof. Let $g_{ijkl} = y_iy_jy_ky_l$. Then this lemma is obtained from Lemma 1.

Lemma 3. *Using a Taylor approximation, we obtain*

$$\hat{\boldsymbol{B}} = \boldsymbol{B} + \boldsymbol{T}^{-1}(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}) + \boldsymbol{R}_0,$$

and

$$\hat{\boldsymbol{B}}^* = \boldsymbol{B} + \boldsymbol{T}^{-1}(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}) + \boldsymbol{R}_0^*,$$

where $\boldsymbol{R}_0$ *and* $\boldsymbol{R}_0^*$ *are the remainder terms of* $\hat{\boldsymbol{B}}$ *and* $\hat{\boldsymbol{B}}^*$ *respectively. If*

Conditions 1 - 4 are satisfied, then

$$E_p\left(\boldsymbol{\alpha}'\boldsymbol{R}_0\right)^2 = O\left(\frac{1}{n^2\lambda_N^4}\right) \quad \textit{and} \quad E_p\left(\boldsymbol{\alpha}'\boldsymbol{R}_0^*\right)^2 = O\left(\frac{1}{n^2\pi_{(K)}^4}\right)$$

uniformly for any given J-vector $\boldsymbol{\alpha}$ with bounded components.

Proof. Let $\hat{\bar{t}}_{jj',\pi}$, $\hat{\bar{t}}_{j0,\pi}$, $\bar{t}_{jj'}$ and $\bar{t}_{j0}$ be the corresponding elements of $\hat{\boldsymbol{T}}$, $\hat{\boldsymbol{t}}$, $\boldsymbol{T}$ and $\boldsymbol{t}$ respectively. Then, using the Taylor linearization for the estimators $\hat{\boldsymbol{B}}$ and $\hat{\boldsymbol{B}}^*$ at the point $(\boldsymbol{T}, \boldsymbol{t})$ respectively, we have

$$\hat{\boldsymbol{B}} = \boldsymbol{B} + \sum_{j=1}^{J}\sum_{j'\le j}\boldsymbol{a}_{jj'}(\hat{\bar{t}}_{jj',\pi} - \bar{t}_{jj'}) + \sum_{j=1}^{J}\boldsymbol{a}_{j0}(\hat{\bar{t}}_{j0,\pi} - \bar{t}_{j0}) + \boldsymbol{R}_0,$$

and

$$\hat{\boldsymbol{B}}^* = \boldsymbol{B} + \sum_{j=1}^{J}\sum_{j'\le j}\boldsymbol{a}_{jj'}(\hat{\bar{t}}^*_{jj',\pi} - \bar{t}_{jj'}) + \sum_{j=1}^{J}\boldsymbol{a}_{j0}(\hat{\bar{t}}^*_{j0,\pi} - \bar{t}_{j0}) + \boldsymbol{R}_0^*,$$

where

$\boldsymbol{a}_{jj'} = \left.\dfrac{\partial\hat{\boldsymbol{B}}}{\partial\hat{\bar{t}}_{jj',\pi}}\right|_{\hat{\boldsymbol{T}}=\boldsymbol{T},\hat{\boldsymbol{t}}=\boldsymbol{t}}$, $\boldsymbol{a}_{j0} = \left.\dfrac{\partial\hat{\boldsymbol{B}}}{\partial\hat{\bar{t}}_{j0,\pi}}\right|_{\hat{\boldsymbol{T}}=\boldsymbol{T},\hat{\boldsymbol{t}}=\boldsymbol{t}}$, and $\boldsymbol{R}_0$ and $\boldsymbol{R}_0^*$ are the remainder terms of $\hat{\boldsymbol{B}}$ and $\hat{\boldsymbol{B}}^*$, respectively. Further, combining the rules of differentiation (see, for example, Graybill (1983)), we obtain the partial derivatives

$$\frac{\partial\hat{\boldsymbol{B}}}{\partial\hat{\bar{t}}_{jj',\pi}} = \left(\frac{\partial\hat{\boldsymbol{T}}^{-1}}{\partial\hat{\bar{t}}_{jj',\pi}}\right)\hat{\boldsymbol{t}} = \big(-\hat{\boldsymbol{T}}^{-1}\boldsymbol{\Lambda}_{jj'}\hat{\boldsymbol{T}}^{-1}\big)\hat{\boldsymbol{t}} = -\hat{\boldsymbol{T}}^{-1}\boldsymbol{\Lambda}_{jj'}\hat{\boldsymbol{B}},$$

where $\boldsymbol{\Lambda}_{jj'}$ is a $J \times J$ matrix with one in positions (j, j') and (j', j), and zeros elsewhere; and

$$\frac{\partial\hat{\boldsymbol{B}}}{\partial\hat{\bar{t}}_{j0,\pi}} = \frac{\partial(\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{t}})}{\partial\hat{\bar{t}}_{j0,\pi}} = \hat{\boldsymbol{T}}^{-1}\left(\frac{\partial\hat{\boldsymbol{t}}}{\partial\hat{\bar{t}}_{j0,\pi}}\right) = \hat{\boldsymbol{T}}^{-1}\boldsymbol{\lambda}_j,$$

where $\boldsymbol{\lambda}_j$ is a J-vector with the j^{th} component being one and zeros elsewhere. Thus

$$\begin{aligned}\hat{\boldsymbol{B}} &= \boldsymbol{B} - \sum_{j=1}^{J}\sum_{j'\leq j} \boldsymbol{T}^{-1}\boldsymbol{\Lambda}_{jj'}\boldsymbol{B}(\hat{\bar{t}}_{jj',\pi} - \bar{t}_{jj'}) + \sum_{j=1}^{J}\boldsymbol{T}^{-1}\boldsymbol{\lambda}_j(\hat{\bar{t}}_{j0,\pi} - \bar{t}_{j0}) + \boldsymbol{R}_0 \\ &= \boldsymbol{B} - \boldsymbol{T}^{-1}(\hat{\boldsymbol{T}} - \boldsymbol{T})\boldsymbol{B} + \boldsymbol{T}^{-1}(\hat{\boldsymbol{t}} - \boldsymbol{t}) + \boldsymbol{R}_0 \\ &= \boldsymbol{B} + \boldsymbol{T}^{-1}(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}) + \boldsymbol{R}_0.\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}\hat{\boldsymbol{B}}^* &= \boldsymbol{B} - \sum_{j=1}^{J}\sum_{j'\leq j} \boldsymbol{T}^{-1}\boldsymbol{\Lambda}_{jj'}\boldsymbol{B}(\hat{\bar{t}}^*_{jj',\pi} - \bar{t}_{jj'}) + \sum_{j=1}^{J}\boldsymbol{T}^{-1}\boldsymbol{\lambda}_j(\hat{\bar{t}}^*_{j0,\pi} - \bar{t}_{j0}) + \boldsymbol{R}^*_0 \\ &= \boldsymbol{B} + \boldsymbol{T}^{-1}(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}) + \boldsymbol{R}^*_0.\end{aligned}$$

Now for any given J-vector $\boldsymbol{\alpha}$, we have

$$\begin{aligned}\boldsymbol{\alpha}'\boldsymbol{R}_0 &= \boldsymbol{\alpha}'\hat{\boldsymbol{B}} - \boldsymbol{\alpha}'\boldsymbol{B} - \boldsymbol{\alpha}'\boldsymbol{T}^{-1}(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}) \\ &= \boldsymbol{\alpha}'\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{t}} - \boldsymbol{\alpha}'\boldsymbol{B} - \boldsymbol{\alpha}'\boldsymbol{T}^{-1}(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}).\end{aligned}$$

Note that the elements of $\hat{\boldsymbol{T}}$ and $\hat{\boldsymbol{t}}$ are the HT estimators, so $(\boldsymbol{\alpha}'\boldsymbol{R}_0)^2$ is a nonlinear function of the HT estimators. By Conditions 1, 3 and Lemma 2, the fourth-order central moment of each HT estimator is $O(n^{-2}\lambda_N^{-4})$. Additionally, the function $(\boldsymbol{\alpha}'\boldsymbol{R}_0)^2$ and its first three derivatives at $(\boldsymbol{T}, \boldsymbol{t})$ evaluate to zero. Therefore, using Theorem 5.4.3 of Fuller (1996) with $\alpha = 1, s = 4$ and $a_N^4 = O(n^{-2}\lambda_N^{-4})$, we see that

$$\mathrm{E}_p\left(\boldsymbol{\alpha}'\boldsymbol{R}_0\right)^2 = O\left(\frac{1}{n^2\lambda_N^4}\right).$$

Similarly, $(\boldsymbol{\alpha}'\boldsymbol{R}_0^*)^2$ is a nonlinear function of IHT estimators. By Conditions 1, 3 and Lemma 2, the fourth-order central moment of each IHT estimator is $O\left(n^{-2}\pi_{(K)}^{-4}\right)$. So using Theorem 5.4.3 of Fuller (1996) with $\alpha = 1, s = 4$ and $a_N^4 = O\left(n^{-2}\pi_{(K)}^{-4}\right)$, we have

$$\mathrm{E}_p\left(\boldsymbol{\alpha}'\boldsymbol{R}_0^*\right)^2 = O\left(\frac{1}{n^2\pi_{(K)}^4}\right).$$

This completes the proof of Lemma 3.

Lemma 4. *If Conditions 1 - 4 are satisfied and* $(n\lambda_N^2)^{-1} = o(1)$, *then*

$$\frac{1}{N}\sum_{i\in U}\left\{E_p\left(\boldsymbol{z}_i'\hat{\boldsymbol{B}}^* - \boldsymbol{z}_i'\boldsymbol{B}\right)^2\right\} = O\left(\frac{1}{n\lambda_N^2}\right)$$

and

$$\frac{1}{K}\sum_{i\in U_2}\left\{E_p\left(\boldsymbol{z}_i'\hat{\boldsymbol{B}}^* - \boldsymbol{z}_i'\boldsymbol{B}\right)^2\right\} = O\left(\frac{1}{n\lambda_N^2}\right).$$

Proof. Note that

$$\max_{j,k\in U}\left|\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})\right| < C$$

and

$$\frac{1}{N}\sum_{k\in U}\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})\right\} = 0.$$

So it follows from Lemma 2 and Cauchy-Schwarz inequality that

$$\mathrm{E}_p\left(\frac{1}{N}\sum_{k\in s}\frac{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})}{\pi_k^*}\right)^2 = O\left(\frac{1}{n\pi_{(K)}^2}\right). \tag{S1.10}$$

Then by Lemma 3 and the fact of $\lambda_N \leq \pi_{(K)}$, we have

$$\begin{aligned} E_p\left(\hat{B}_j^* - B_j\right)^2 &= E_p\left\{\sum_{k\in s}\frac{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})}{N\pi_k^*} + \boldsymbol{\lambda}_j'\boldsymbol{R}_0^*\right\}^2 \\ &\leq 2E_p\left\{\sum_{k\in s}\frac{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})}{N\pi_k^*}\right\}^2 + 2E_p\left(\boldsymbol{\lambda}_j'\boldsymbol{R}_0^*\right)^2 \\ &= O\left(\frac{1}{n\lambda_N^2} + \frac{1}{n^2\lambda_N^4}\right). \qquad \text{(S1.11)} \end{aligned}$$

Thus using Condition 3 and $(n\lambda_N^2)^{-1} = o(1)$, we obtain

$$\begin{aligned} \frac{1}{N}\sum_{i\in U}\left\{\mathrm{E}_p\left(\boldsymbol{z}_i'\hat{\boldsymbol{B}}^* - \boldsymbol{z}_i'\boldsymbol{B}\right)^2\right\} &\leq \mathrm{E}_p\left\|\hat{\boldsymbol{B}}^* - \boldsymbol{B}\right\|^2\cdot\frac{1}{N}\sum_{i\in U}\|\boldsymbol{z}_i\|^2 \\ &= \sum_{j=1}^{J}\left\{\mathrm{E}_p\left(\hat{B}_j^* - B_j\right)^2\right\}\cdot\frac{1}{N}\sum_{i\in U}\|\boldsymbol{z}_i\|^2 \\ &= O\left(\frac{1}{n\lambda_N^2}\right). \end{aligned}$$

Similarly, we can prove that

$$\frac{1}{K}\sum_{i\in U_2}\left\{\mathrm{E}_p\left(\boldsymbol{z}_i'\hat{\boldsymbol{B}}^* - \boldsymbol{z}_i'\boldsymbol{B}\right)^2\right\} = O\left(\frac{1}{n\lambda_N^2}\right).$$

This completes the proof of Lemma 4.

Lemma 5. *If Conditions 2 and 4 hold and $\sqrt{n}K/N = O(1)$, then*

$$E_p\left|\frac{1}{N^2}\sum_{i,j\in U}\left[\Delta_{ij}^*(I_iI_j - \pi_{ij})\right]\right| = O\left(n^{-1-\kappa}\right),$$

where $\kappa = \min\left\{\frac{\alpha}{2}, \frac{1}{4}\right\}$.

Proof. Note that

$$\mathrm{E}_p\left(\frac{1}{N^2}\sum_{i,j\in U}\left[\Delta_{ij}^*(I_iI_j - \pi_{ij})\right]\right)^2$$

$$
\begin{aligned}
&= \frac{1}{N^4} \sum_{i,j,k,l \in U} \left\{ \Delta_{ij}^* \Delta_{kl}^* \cdot \mathrm{E}_p \left[(I_i I_j - \pi_{ij})(I_k I_l - \pi_{kl}) \right] \right\} \\
&= \frac{1}{N^4} \sum_{i,k \in U} \left\{ \Delta_{ii}^* \Delta_{kk}^* \cdot \mathrm{E}_p \left[(I_i - \pi_i)(I_k - \pi_k) \right] \right\} \\
&\quad + \frac{2}{N^4} \sum_{i \in U, k \neq l \in U} \left\{ \Delta_{ii}^* \Delta_{kl}^* \cdot \mathrm{E}_p \left[(I_i - \pi_i)(I_k I_l - \pi_{kl}) \right] \right\} \\
&\quad + \frac{1}{N^4} \sum_{i \neq j \in U, k \neq l \in U} \left\{ \Delta_{ij}^* \Delta_{kl}^* \cdot \mathrm{E}_p \left[(I_i I_j - \pi_{ij})(I_k I_l - \pi_{kl}) \right] \right\} \\
&\triangleq A_{1N} + A_{2N} + A_{3N}.
\end{aligned}
$$

To finish the proof of Lemma 5, we need only to prove that the orders of A_{1N}, A_{2N} and A_{3N} are all $O\left(n^{-2-2\kappa}\right)$.

For A_{1N}, by Condition 2 and $\max_{i,j \in U} |\Delta_{ij}^*| \leq C$, we have

$$
\begin{aligned}
|A_{1N}| &\leq \frac{1}{N^4} \sum_{i,k \in U} \left| \Delta_{ii}^* \Delta_{kk}^* \cdot \mathrm{E}_p \left[(I_i - \pi_i)(I_k - \pi_k) \right] \right| \\
&\leq C N^{-2} \max_{k \neq l \in U} |\pi_{ik} - \pi_i \pi_k| + C N^{-3} \\
&= O(n^{-3}).
\end{aligned}
$$

Next, we consider A_{2N}. Note that $\Delta_{kl}^* = \pi_{kl} - \pi_k \pi_l$ for $k \in U_1$ and $l \in U$. Then, it follows from Condition 2 that

$$
\begin{aligned}
\sum_{k \neq l \in U} |\Delta_{kl}^*| &= \sum_{k \neq l \in U_1} |\pi_{kl} - \pi_k \pi_l| + 2 \sum_{k \in U_1, l \in U_2} |\pi_{kl} - \pi_k \pi_l| + \sum_{k \neq l \in U_2} |\Delta_{kl}^*| \\
&= O\left(N^2 n^{-1} + K^2\right). \qquad \text{(S1.12)}
\end{aligned}
$$

Thus by Condition 4 and $\sqrt{n}K/N = O(1)$, we obtain

$$
\begin{aligned}
|A_{2N}| &\leq \frac{2}{N^4} \sum_{i\in U, k\neq l \in U} \left| \Delta_{ii}^* \Delta_{kl}^* \cdot \mathrm{E}_p \left[(I_i - \pi_i)(I_k I_l - \pi_{kl}) \right] \right| \\
&\leq \max_{i,k,l \in D_{3,N}} |\pi_{ikl} - \pi_i \pi_{kl}| \cdot \frac{C}{N^3} \sum_{k\neq l\in U} |\Delta_{kl}^*| + \frac{C}{N^4} \sum_{k\neq l \in U} |\Delta_{kl}^*| \\
&\leq \frac{C}{nN} \cdot \left(\frac{1}{n} + \frac{K^2}{N^2} \right) \\
&= O(n^{-3}).
\end{aligned}
$$

Finally, we consider A_{3N}. From Condition 2, we observe that

$$
\begin{aligned}
\sum_{(j,k,l)\in D_{3,N}} |\Delta_{kj}^* \Delta_{kl}^*| &= \sum_{\substack{k\in U_1,\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}\Delta_{kl}| + \sum_{\substack{k\in U_2,\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}^* \Delta_{kl}^*| \\
&= \sum_{\substack{k\in U_1,\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}\Delta_{kl}| + \sum_{\substack{k\in U_2, j\neq l\in U_1\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}\Delta_{kl}| \\
&\quad + 2 \sum_{\substack{k\in U_2, j\in U_1, l\in U_2\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}\Delta_{kl}^*| + \sum_{\substack{k\in U_2, j\neq l\in U_2,\\ (j,k,l)\in D_{3,N}}} |\Delta_{kj}^*\Delta_{kl}^*| \\
&= O\left(N^3 n^{-2} + N^2 K n^{-2} + N K^2 n^{-1} + K^3 \right) \\
&= O\left(N^3 n^{-2} + N K^2 n^{-1} + K^3 \right) \qquad \text{(S1.13)}
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{k\neq l\in U} |\Delta_{kl}^{*2}| &= \sum_{k\neq l\in U_1} |\Delta_{kl}^2| + \sum_{k\neq l\in U_2} |\Delta_{kl}^{*2}| + 2 \sum_{k\in U_1, l\in U_2} |\Delta_{kl}^2| \\
&= O\left(N^2 n^{-2} + K^2 \right). \qquad \text{(S1.14)}
\end{aligned}
$$

Let $\tilde{D}_{4,N} = \{(i,j,k,l) | i \neq j \in U, k \neq l \in U \text{ and } (i,j,k,l) \notin D_{4,N}\}$. Then combining Equations (S1.12) - (S1.14), Condition 4 and $\sqrt{n}K/N = O(1)$, we have

$$\begin{aligned}
|A_{3N}| &\le \frac{1}{N^4} \sum_{i\neq j\in U, k\neq l\in U} \left|\Delta^*_{ij}\Delta^*_{kl}\cdot(\pi_{ijkl}-\pi_{ij}\pi_{kl})\right| \\
&\le \frac{1}{N^4} \sum_{(i,j,k,l)\in \tilde{D}_{4,N}} \left|\Delta^*_{ij}\Delta^*_{kl}\cdot(\pi_{ijkl}-\pi_{ij}\pi_{kl})\right| \\
&\quad + \max_{(i,j,k,l)\in D_{4,N}} |\pi_{ijkl}-\pi_{ij}\pi_{kl}| \cdot \left(\frac{1}{N^2}\sum_{i\neq j\in U}|\Delta^*_{ij}|\right)^2 \\
&\le \frac{2}{N^4}\sum_{(j,k,l)\in D_{3,N}} \left|\Delta^*_{kj}\Delta^*_{kl}\right| + \frac{2}{N^4}\sum_{k\neq l\in U}\left|\Delta^{*2}_{kl}\right| \\
&\quad + \frac{1}{n^\alpha}\cdot\left(\frac{1}{N^2}\sum_{i\neq j\in U}|\Delta^*_{ij}|\right)^2 \\
&\le C\cdot\left(n^{-5/2}+n^{-3}+n^{-2-\alpha}\right) \\
&= O\left(n^{-2-2\kappa}\right),
\end{aligned}$$

where $\kappa = \min\left\{\frac{\alpha}{2}, \frac{1}{4}\right\}$. This completes the proof of Lemma 5.

Lemma 6. *Let g_i be a random variable. If Condition 2 holds and $\sqrt{n}K/N = O(1)$, then we have*

$$E_p\left|\frac{1}{N^2}\sum_{i,j\in U}\left(g_i\Delta^*_{ij}\right)\right| \le Cn^{-1}\cdot\left(\frac{1}{N}\sum_{i\in U}E_p|g_i| + \frac{1}{K}\sum_{i\in U_2}E_p|g_i|\right).$$

Proof. Observe that $|\Delta^*_{ij}| = |\pi_{ij}-\pi_i\pi_j|$ for $i\in U_1$ and $j\in U$. Using Condition 2 and $\sqrt{n}K/N = O(1)$, we have

$$\mathrm{E}_p\left|\frac{1}{N^2}\sum_{i,j\in U}\left(g_i\Delta^*_{ij}\right)\right| \le \frac{1}{N^2}\sum_{i,j\in U}\left(|\Delta^*_{ij}|\cdot \mathrm{E}_p|g_i|\right)$$

$$
\begin{aligned}
&= \frac{1}{N^2} \sum_{i\in U_1, j\in U_1} \Big(|\pi_{ij} - \pi_i\pi_j| \cdot \mathrm{E}_p|g_i| \Big) \\
&\quad + \frac{K}{N}\frac{1}{NK} \sum_{i\in U_1, j\in U_2} \Big(|\pi_{ij} - \pi_i\pi_j| \cdot \mathrm{E}_p|g_i| \Big) \\
&\quad + \frac{K}{N}\frac{1}{NK} \sum_{i\in U_2, j\in U_1} \Big(|\pi_{ij} - \pi_i\pi_j| \cdot \mathrm{E}_p|g_i| \Big) \\
&\quad + \frac{K^2}{N^2}\frac{1}{K^2} \sum_{i\in U_2, j\in U_2} \Big(|\Delta_{ij}^*| \cdot \mathrm{E}_p|g_i| \Big) \\
&\le Cn^{-1} \cdot \left(\frac{1}{N} \sum_{i\in U} \mathrm{E}_p|g_i| + \frac{1}{K} \sum_{i\in U_2} \mathrm{E}_p|g_i| \right).
\end{aligned}
$$

This completes the proof of Lemma 6.

Lemma 7. *If Conditions 1 - 4 hold, then*

$$
E_p(b_N^2) \triangleq E_p \left\{ \frac{1}{N} \sum_{k\in U} \frac{(\pi_k^* - I_k)(\boldsymbol{z}_k'\hat{\boldsymbol{B}}^* - \boldsymbol{z}_k'\boldsymbol{B})}{\pi_k^*} \right\}^2 = O\left(\frac{1}{n^2\lambda_N^6} \right).
$$

Proof. Write $C_{kl} = \boldsymbol{z}_k'\boldsymbol{T}^{-1}\left(\boldsymbol{z}_l y_l - \boldsymbol{z}_l\boldsymbol{z}_l'\boldsymbol{B}\right)$. Then

$$
\begin{aligned}
\mathrm{E}_p(b_N^2) &= \mathrm{E}_p \left\{ \frac{1}{N} \sum_{k\in U} \frac{(\pi_k^* - I_k)\left(\boldsymbol{z}_k'\boldsymbol{T}^{-1}(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}) + \boldsymbol{z}_k'\boldsymbol{R}_0^* \right)}{\pi_k^*} \right\}^2 \\
&= \mathrm{E}_p \left\{ -\frac{1}{N^2} \sum_{k,l\in U} C_{kl}\Phi_k^*\Phi_l^* + \frac{1}{N} \sum_{k\in U} \frac{\boldsymbol{z}_k'\boldsymbol{R}_0^*(\pi_k^* - I_k)}{\pi_k^*} \right\}^2 \\
&\le 2\mathrm{E}_p \left\{ \frac{1}{N^2} \sum_{k,l\in U} C_{kl}\Phi_k^*\Phi_l^* \right\}^2 + 2\mathrm{E}_p \left\{ \frac{1}{N} \sum_{k\in U} \frac{\boldsymbol{z}_k'\boldsymbol{R}_0^*(\pi_k^* - I_k)}{\pi_k^*} \right\}^2 \\
&\triangleq b_{1N} + b_{2N}.
\end{aligned}
$$

Condition 3 implies $\max_{k,l\in U} |C_{kl}| \le C$. Thus, by Lemma 1, $b_{1N} = O(n^{-2}\lambda_N^{-4})$.

Further, using Lemma 3 and Cauchy-Schwarz inequality, we have

$$b_{2N} \leq \frac{C}{N}\mathrm{E}_p\left\{\sum_{k\in U}\frac{(\boldsymbol{z}_k'\boldsymbol{R}_0^*)^2(\pi_k^*-I_k)^2}{\pi_k^{*2}}\right\} \leq \frac{C}{N\lambda_N^2}\sum_{k\in U}\mathrm{E}_p(\boldsymbol{z}_k'\boldsymbol{R}_0^*)^2 = O(n^{-2}\lambda_N^{-6}).$$

Hence $\mathrm{E}_p(b_N^2) \leq b_{1N} + b_{2N} = O(n^{-2}\lambda_N^{-6})$. This completes the proof of Lemma 7.

S1.4 Proof of Theorem 3

Proof. Using (S1.11) and Cauchy-Schwarz inequality, we have

$$\mathrm{E}_p\left|\hat{B}_j^* - B_j\right| \leq \left\{\mathrm{E}_p\left(\hat{B}_j^* - B_j\right)^2\right\}^{1/2} = O\left(n^{-1/2}\lambda_N^{-1}\right).$$

Thus, the improved estimator $\hat{\boldsymbol{B}}^*$ is asymptotically design unbiased and design consistent by the condition $(n\lambda_N^2)^{-1} = o(1)$.

Next, we evaluate the design variance-covariance matrix of $\hat{\boldsymbol{B}}^*$. Using Lemma 3, it is readily seen that

$$\begin{aligned}
\mathrm{MSE}_p(\hat{\boldsymbol{B}}^*) &= \mathrm{E}_p\left\{(\hat{\boldsymbol{B}}^* - \boldsymbol{B})(\hat{\boldsymbol{B}}^* - \boldsymbol{B})'\right\} \\
&= \mathrm{E}_p\left\{\left(\boldsymbol{T}^{-1}(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}) + \boldsymbol{R}_0^*\right)\left(\boldsymbol{T}^{-1}(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}) + \boldsymbol{R}_0^*\right)'\right\} \\
&= \boldsymbol{T}^{-1}\mathrm{E}_p\left\{\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)'\right\}\boldsymbol{T}^{-1} + \mathrm{E}_p\left(\boldsymbol{R}_0^*\boldsymbol{R}_0^{*'}\right) \\
&\quad + \mathrm{E}_p\left\{\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)\boldsymbol{R}_0^{*'}\right\} + \mathrm{E}_p\left\{\boldsymbol{R}_0^*\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)'\boldsymbol{T}^{-1}\right\}.
\end{aligned}$$

Note that the second term $\mathrm{E}_p\left(\boldsymbol{R}_0^*\boldsymbol{R}_0^{*'}\right)$ is a $J\times J$ matrix with $\mathrm{E}_p\left(\boldsymbol{\lambda}_j'\boldsymbol{R}_0^*\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^*\right)$ in position (j, j'). By Lemma 3, we have

$$\mathrm{E}_p\left|\boldsymbol{\lambda}_j'\boldsymbol{R}_0^*\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^*\right| \leq \sqrt{\mathrm{E}_p\left(\boldsymbol{\lambda}_j'\boldsymbol{R}_0^*\right)^2\mathrm{E}_p\left(\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^*\right)^2} = O\left(n^{-2}\pi_{(K)}^{-4}\right).$$

It is clear that the element of the third term $\mathrm{E}_p\left\{\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\boldsymbol{R}_0^{*\prime}\right\}$ in position (j,j') is $\mathrm{E}_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^{*}\right\}$. Combining (S1.10) and Lemma 3, we obtain

$$\begin{aligned}\mathrm{E}_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^{*}\right\} &\leq \left\{\mathrm{E}_p\left[\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\right]^2\mathrm{E}_p\left(\boldsymbol{\lambda}_{j'}'\boldsymbol{R}_0^{*}\right)^2\right\}^{1/2}\\ &= O\left(n^{-3/2}\pi_{(K)}^{-3}\right).\end{aligned}$$

Since

$$\left\{\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\boldsymbol{R}_0^{*\prime}\right\}'=\left\{\boldsymbol{R}_0^{*}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)'\boldsymbol{T}^{-1}\right\},$$

the order of each element in the last term $\mathrm{E}_p\left\{\boldsymbol{R}_0^{*}\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)'\boldsymbol{T}^{-1}\right\}$ is $O\left(n^{-3/2}\pi_{(K)}^{-3}\right)$ too. Therefore,

$$\begin{aligned}&\mathrm{MSE}_p(\hat{\boldsymbol{B}}^{*})\\ &=\boldsymbol{T}^{-1}\mathrm{E}_p\left\{\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)'\right\}\boldsymbol{T}^{-1}+O\left(n^{-3/2}\lambda_N^{-3}\right).\end{aligned}\quad\text{(S1.15)}$$

Similarly, we can prove that

$$\begin{aligned}&\mathrm{MSE}_p(\hat{\boldsymbol{B}})\\ &=\boldsymbol{T}^{-1}\mathrm{E}_p\left\{\left(\hat{\boldsymbol{t}}-\hat{\boldsymbol{T}}\boldsymbol{B}\right)\left(\hat{\boldsymbol{t}}-\hat{\boldsymbol{T}}\boldsymbol{B}\right)'\right\}\boldsymbol{T}^{-1}+O\left(n^{-3/2}\lambda_N^{-3}\right).\end{aligned}\quad\text{(S1.16)}$$

Let $E_k=y_k-\boldsymbol{z}_k'\boldsymbol{B}, k\in U$. For the element of $\mathrm{E}_p\left\{\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)'\right\}$ in position (j,j'), denoted as $v_{jj'}^{*}$, we have

$$v_{jj'}^{*}=\mathrm{E}_p\left\{\boldsymbol{\lambda}_j'\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\boldsymbol{\lambda}_{j'}'\left(\hat{\boldsymbol{t}}^{*}-\hat{\boldsymbol{T}}^{*}\boldsymbol{B}\right)\right\}$$

$$= \mathrm{E}_p \left\{ \frac{1}{N^2} \sum_{k,l \in s} \frac{\left\{ \boldsymbol{\lambda}_j' \boldsymbol{z}_k \left(y_k - \boldsymbol{z}_k' \boldsymbol{B} \right) \right\} \left\{ \boldsymbol{\lambda}_{j'}' \boldsymbol{z}_l \left(y_l - \boldsymbol{z}_l' \boldsymbol{B} \right) \right\}}{\pi_k^* \pi_l^*} \right\}$$

$$= \mathrm{E}_p \left\{ \frac{1}{N^2} \sum_{k,l \in U} \frac{\left(z_{jk} E_k \right) \left(z_{j'l} E_l \right)}{\pi_k^* \pi_l^*} I_k I_l \right\}$$

$$= \frac{1}{N^2} \sum_{k,l \in U} \frac{\pi_{kl}}{\pi_k^* \pi_l^*} \left(z_{jk} E_k \right) \left(z_{j'l} E_l \right).$$

Then, by the fact that $\sum_U z_{jk} E_k = 0$, we see that

$$v_{jj'}^* = \frac{1}{N^2} \sum_{k,l \in U} \frac{\pi_{kl} - \pi_k \pi_l^* - \pi_k^* \pi_l + \pi_k^* \pi_l^*}{\pi_k^* \pi_l^*} \left(z_{jk} E_k \right) \left(z_{j'l} E_l \right)$$

$$= \frac{1}{N^2} \sum_{k,l \in U} \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \left(z_{jk} E_k \right) \left(z_{j'l} E_l \right),$$

where $\Delta_{kl}^* = \pi_{kl} - \pi_k \pi_l^* - \pi_k^* \pi_l + \pi_k^* \pi_l^*$. Thus, from (S1.15), the design variance-covariance matrix of $\hat{\boldsymbol{B}}^*$ is obtained as claimed in Theorem 3.

S1.5 Proof of Theorem 4

Proof. Define

$$A_N(j,j') = \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ z_{jk} z_{j'l} \left(y_k - \boldsymbol{z}_k' \boldsymbol{B} \right) \left(y_l - \boldsymbol{z}_l' \boldsymbol{B} \right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l - \pi_{kl}}{\pi_{kl}} \right\} \right|$$

and

$$B_N(j,j') = \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ z_{jk} z_{j'l} \left(y_k - \boldsymbol{z}_k' \boldsymbol{B} \right) \left(\boldsymbol{z}_l' \boldsymbol{B} - \boldsymbol{z}_l' \hat{\boldsymbol{B}}^* \right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l}{\pi_{kl}} \right\} \right|$$

$$+ \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ z_{jk} z_{j'l} \left(\boldsymbol{z}_k' \boldsymbol{B} - \boldsymbol{z}_k' \hat{\boldsymbol{B}}^* \right) \left(y_l - \boldsymbol{z}_l' \boldsymbol{B} \right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l}{\pi_{kl}} \right\} \right|$$

$$+ \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ z_{jk} z_{j'l} \left(\boldsymbol{z}_k' \boldsymbol{B} - \boldsymbol{z}_k' \hat{\boldsymbol{B}}^* \right) \left(\boldsymbol{z}_l' \boldsymbol{B} - \boldsymbol{z}_l' \hat{\boldsymbol{B}}^* \right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l}{\pi_{kl}} \right\} \right|.$$

Then

$$\mathrm{E}_p \left|\hat{v}^*_{jj'} - v^*_{jj'}\right| \leq A_N(j,j') + B_N(j,j').$$

Note that, for any fixed (j, j'),

$$\max_{k,l\in U}\left|\frac{z_{jk}z_{j'l}\left(y_k - \boldsymbol{z}_k'\boldsymbol{B}\right)\left(y_l - \boldsymbol{z}_l'\boldsymbol{B}\right)}{\pi^*_k\pi^*_l\pi_{kl}}\right| \leq \frac{C}{\lambda_N^2\lambda_N^*}.$$

From Definition 1, we have $K \leq \lambda_N^{-1}$. Then using the condition $(n\lambda_N^3)^{-1} = o(1)$, we see that $\sqrt{n}K/N = O(1)$. Thus, from Lemma 5, it follows that $A_N(j,j') = O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right)$, where $\kappa = \min\left\{\frac{\alpha}{2}, \frac{1}{4}\right\}$.

Next, we prove $B_N(j,j') = O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$. By Conditions 1 - 3, it is seen that

$$\max_{k,l\in U}\left|\frac{z_{jk}z_{j'l}I_kI_l}{\pi^*_k\pi^*_l\pi_{kl}}\left(y_k - \boldsymbol{z}_{\boldsymbol{k}}'\boldsymbol{B}\right)\right| \leq \frac{C}{\lambda_N^2\lambda_N^*}.$$

Then using Lemmas 4 and 6 and the fact of $\sqrt{n}K/N = O(1)$, the order of the first term in $B_N(j,j')$ is $O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$. Similarly, it is straightforward to show that the orders of the other terms in $B_N(j,j')$ are $O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$ too.

Thus, by the condition $(n\lambda_N^3)^{-1} = o(1)$, we obtain

$$\mathrm{E}_p \left|\hat{v}^*_{jj'} - v^*_{jj'}\right| \leq A_N(j,j') + B_N(j,j') = O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right).$$

This completes the proof of Theorem 4.

S1.6 Proof of Theorem 5

Following (S1.15) and (S1.16), we have

$$tr\left[\mathrm{MSE}_p(\hat{\boldsymbol{B}}^*)\right] = \sum_{j=1}^{J} E_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)\right\}^2 + O\left(n^{-3/2}\lambda_N^{-3}\right)$$

and

$$tr\left[\mathrm{MSE}_p(\hat{\boldsymbol{B}})\right] = \sum_{j=1}^{J} E_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}\right)\right\}^2 + O\left(n^{-3/2}\lambda_N^{-3}\right).$$

Note that $\boldsymbol{T}^{-1}\left(\boldsymbol{t} - \boldsymbol{T}\boldsymbol{B}\right) = \boldsymbol{0}$. By Conditions 1 and 2, Theorem 2 and $(n\lambda_N^6)^{-1} = o(1)$, we observe that, for any j,

$$\begin{aligned}
&E_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}}^* - \hat{\boldsymbol{T}}^*\boldsymbol{B}\right)\right\}^2 \\
&= E_p\left\{\sum_{k\in s}\frac{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})}{\pi_k^*} - \sum_{k\in U}\left[\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})\right]\right\}^2 \\
&\le E_p\left\{\sum_{k\in s}\frac{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})}{\pi_k} - \sum_{k\in U}\left[\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}(\boldsymbol{z}_k y_k - \boldsymbol{z}_k\boldsymbol{z}_k'\boldsymbol{B})\right]\right\}^2 + o\left(n^{-1}\right) \\
&= E_p\left\{\boldsymbol{\lambda}_j'\boldsymbol{T}^{-1}\left(\hat{\boldsymbol{t}} - \hat{\boldsymbol{T}}\boldsymbol{B}\right)\right\}^2 + o\left(n^{-1}\right).
\end{aligned}$$

Thus

$$tr\left[\mathrm{MSE}_p(\hat{\boldsymbol{B}}^*)\right] \le tr\left[\mathrm{MSE}_p(\hat{\boldsymbol{B}})\right] + o\left(n^{-1}\right).$$

This completes the proof of Theorem 5.

S1.7 Proof of Theorem 6

We first show that the improved estimator $\hat{\hat{t}}_{yr}^*$ is asymptotically design unbiased and design consistent. By Markov's inequality, it suffices to prove that

$$\mathrm{E}_p \left| \hat{\bar{t}}^*_{yr} - \bar{t}_y \right| = o(1).$$

Note that

$$\hat{\bar{t}}^*_{yr} - \bar{t}_y = \frac{1}{N} \sum_{k \in U} \frac{(I_k - \pi^*_k)(y_k - \boldsymbol{z_k'B})}{\pi^*_k} + \frac{1}{N} \sum_{k \in U} \frac{(\pi^*_k - I_k)(\boldsymbol{z_k'}\hat{\boldsymbol{B}}^* - \boldsymbol{z_k'B})}{\pi^*_k}$$

$$\triangleq a_N + b_N.$$

Then

$$\mathrm{E}_p \left| \hat{\bar{t}}^*_{yr} - \bar{t}_y \right| \leq \mathrm{E}_p \left| a_N \right| + \mathrm{E}_p \left| b_N \right|.$$

By Conditions 1 - 4 and Lemma 2, we have

$$\begin{aligned}
\mathrm{E}_p(a_N^2) &= \mathrm{E}_p \left\{ \frac{1}{N} \sum_{k \in U} \frac{(I_k - \pi^*_k)(y_k - \boldsymbol{z_k'B})}{\pi^*_k} \right\}^2 \\
&= \frac{1}{N^2} \sum_{k,l \in U} \left\{ \frac{\pi_{kl} - \pi_k \pi^*_l - \pi^*_k \pi_l + \pi^*_k \pi^*_l}{\pi^*_k \pi^*_l} (y_k - \boldsymbol{z_k'B})(y_l - \boldsymbol{z_l'B}) \right\} \\
&= O(n^{-1} \lambda_N^{-2}). \qquad \text{(S1.17)}
\end{aligned}$$

Further, from Lemma 7, we see that $\mathrm{E}_p\left(b_N^2\right) = O(n^{-2}\lambda_N^{-6})$. Thus, using Cauchy-Schwarz inequality and the condition $(n\lambda_N^4)^{-1} = o(1)$, we obtain

$$\mathrm{E}_p \left| \hat{\bar{t}}^*_{yr} - \bar{t}_y \right| \leq \mathrm{E}_p \left| a_N \right| + \mathrm{E}_p \left| b_N \right| = O(n^{-1/2}\lambda_N^{-1}) = o(1).$$

Next, we turn to evaluate the design mean squared error of $\hat{\bar{t}}^*_{yr}$. Observe that,

$$\mathrm{MSE}_p\left(\hat{\bar{t}}^*_{yr}\right) = \mathrm{E}_p\left(\hat{\bar{t}}^*_{yr} - \bar{t}_y\right)^2 = \mathrm{E}_p(a_N^2) + \mathrm{E}_p(b_N^2) + 2\mathrm{E}_p(a_N b_N).$$

Again, by Lemma 7, we see that $\mathrm{E}_p(b_N^2) = O(n^{-2}\lambda_N^{-6})$. Further, combining (S1.17) and Cauchy-Schwarz inequality, it is obvious that

$$\mathrm{E}_p|a_N b_N| \le \left\{\mathrm{E}_p(a_N^2)\mathrm{E}_p(b_N^2)\right\}^{1/2} = O(n^{-3/2}\lambda_N^{-4}).$$

Hence, from the condition $(n\lambda_N^4)^{-1} = o(1)$, the design mean squared error of $\hat{\hat{t}}_{yr}^*$ is

$$\begin{aligned}\mathrm{MSE}_p\left(\hat{\hat{t}}_{yr}^*\right) &= \mathrm{E}_p(a_N^2) + O(n^{-2}\lambda_N^{-6}) + O(n^{-3/2}\lambda_N^{-4}) \\ &= \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}E_k E_l + O(n^{-3/2}\lambda_N^{-4}). \qquad \text{(S1.18)}\end{aligned}$$

Finally, to prove that $\widehat{\mathrm{AMSE}}_p\left(\hat{\hat{t}}_{yr}^*\right)$ is an asymptotically design unbiased estimator, it suffices to show that

$$\mathrm{E}_p\left|\widehat{\mathrm{AMSE}}_p\left(\hat{\hat{t}}_{yr}^*\right) - \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}E_k E_l\right| = O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right).$$

Define

$$A_N = \mathrm{E}_p\left|\frac{1}{N^2}\sum_{k,l\in U}\left\{(y_k - \boldsymbol{z}_{\boldsymbol{k}}'\boldsymbol{B})(y_l - \boldsymbol{z}_{\boldsymbol{l}}'\boldsymbol{B})\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}\frac{I_k I_l - \pi_{kl}}{\pi_{kl}}\right\}\right|$$

and

$$\begin{aligned}B_N = \mathrm{E}_p&\left|\frac{1}{N^2}\sum_{k,l\in U}\left\{2(y_k - \boldsymbol{z}_{\boldsymbol{k}}'\boldsymbol{B})\left(\boldsymbol{z}_{\boldsymbol{l}}'\boldsymbol{B} - \boldsymbol{z}_{\boldsymbol{l}}'\hat{\boldsymbol{B}}^*\right)\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}\frac{I_k I_l}{\pi_{kl}}\right\}\right| \\ &+ \mathrm{E}_p\left|\frac{1}{N^2}\sum_{k,l\in U}\left\{\left(\boldsymbol{z}_{\boldsymbol{k}}'\boldsymbol{B} - \boldsymbol{z}_{\boldsymbol{k}}'\hat{\boldsymbol{B}}^*\right)\left(\boldsymbol{z}_{\boldsymbol{l}}'\boldsymbol{B} - \boldsymbol{z}_{\boldsymbol{l}}'\hat{\boldsymbol{B}}^*\right)\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}\frac{I_k I_l}{\pi_{kl}}\right\}\right|.\end{aligned}$$

Then

$$\mathrm{E}_p\left|\widehat{\mathrm{AMSE}}_p\left(\hat{\hat{t}}_{yr}^*\right) - \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}E_k E_l\right| \le A_N + B_N.$$

Note that

$$\max_{k,l\in U}\left|\frac{(y_k-\boldsymbol{z}_k'\boldsymbol{B})(y_l-\boldsymbol{z}_l'\boldsymbol{B})}{\pi_k^*\pi_l^*\pi_{kl}}\right|\leq\frac{C}{\lambda_N^2\lambda_N^*}.$$

By Definition 1 and the condition $(n\lambda_N^8)^{-1}=o(1)$, we have $\sqrt{n}K/N=O(1)$. Thus, from Lemma 5, it follows that $A_N=O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right)$, where $\kappa=\min\left\{\frac{\alpha}{2},\frac{1}{4}\right\}$.

Next, we prove $B_N=O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$. By Conditions 1 - 3, it is seen that

$$\max_{k,l\in U}\left|\frac{2I_kI_l}{\pi_k^*\pi_l^*\pi_{kl}}(y_k-\boldsymbol{z}_{\boldsymbol{k}}'\boldsymbol{B})\right|\leq\frac{C}{\lambda_N^2\lambda_N^*}.$$

Then using Lemmas 4 and 6 and the fact of $\sqrt{n}K/N=O(1)$, the order of the first term in B_N is $O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$. Similarly, it is straightforward to show that the orders of the other terms in B_N are $O\left(n^{-2}\lambda_N^{-4}\lambda_N^{*-1}\right)$ too.

Thus, by the condition $(n\lambda_N^8)^{-1}=o(1)$, we obtain

$$\mathrm{E}_p\left|\widehat{\mathrm{AMSE}}_p\left(\hat{\hat{t}}_{yr}^*\right)-\frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}E_kE_l\right|\leq A_N+B_N=O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right).$$

This completes the proof of Theorem 6.

S1.8 Proof of Theorem 7

Proof. From Särndal et al. (1992), the design mean squared error of the estimator $\hat{\hat{t}}_{yr}$ is

$$\mathrm{MSE}_p\left(\hat{\hat{t}}_{yr}\right)=\frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}}{\pi_k\pi_l}E_kE_l+o(n^{-1}),\tag{S1.19}$$

where $\Delta_{kl} = \pi_{kl} - \pi_k\pi_l$. By Conditions 1 and 2, Theorem 2 and the condition $(n\lambda_N^8)^{-1} = o(1)$, we observe that

$$\begin{aligned}\frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}E_kE_l &= E_p\left(\frac{1}{N}\sum_{k\in s}\frac{E_k}{\pi_k^*} - \frac{1}{N}\sum_{k\in U}E_k\right)^2 \\ &\leq E_p\left(\frac{1}{N}\sum_{k\in s}\frac{E_k}{\pi_k} - \frac{1}{N}\sum_{k\in U}E_k\right)^2 + o(n^{-1}) \\ &= \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}}{\pi_k\pi_l}E_kE_l + o(n^{-1}).\end{aligned}$$

Thus, combining (S1.18), (S1.19) and the condition $(n\lambda_N^8)^{-1} = o(1)$, we obtain that

$$\text{MSE}_p\left(\hat{\hat{t}}_{yr}^*\right) \leq \text{MSE}_p\left(\hat{\hat{t}}_{yr}\right) + o(n^{-1}).$$

This completes the proof of Theorem 7.

S1.9 Lemmas for the proofs of Theorems 8 - 9

First, we introduce some notations. Following Breidt and Opsomer (2000), we have $m_i = f(\bar{\boldsymbol{t}}_i, 0)$ and $\hat{m}_i = f(\hat{\bar{\boldsymbol{t}}}_i, \delta)$, where

$$\bar{\boldsymbol{t}}_i = [\bar{t}_{ig}]_{g=1}^G = \left[\frac{1}{N}\sum_{k\in U}\frac{1}{h_N}\mathcal{K}\left(\frac{x_j - x_i}{h_N}\right)z_{igk}^\dagger\right]_{g=1}^G \triangleq \left[\frac{1}{N}\sum_{k\in U}z_{igk}^*\right]_{g=1}^G,$$

and its HT estimator

$$\hat{\bar{\boldsymbol{t}}}_i = \left[\hat{\bar{t}}_{ig}\right]_{g=1}^G = \left[\frac{1}{N}\sum_{k\in s}\frac{1}{h_N}\mathcal{K}\left(\frac{x_j - x_i}{h_N}\right)\frac{z_{igk}^\dagger}{\pi_k}\right]_{g=1}^G \triangleq \left[\frac{1}{N}\sum_{k\in s}\frac{z_{igk}^*}{\pi_k}\right]_{g=1}^G$$

with $G = 3q + 2$,

$$z_{igk}^{\dagger} = \begin{cases} (x_k - x_i)^{g-1}, & g \le G_1, \\ (x_k - x_i)^{g-G_1-1} y_k, & g > G_1, \end{cases}$$

and $G_1 = 2q + 1$.

Similarly, we can write $\hat{m}_i^* = f(\hat{\bar{\boldsymbol{t}}}_i^*, \delta)$ with

$$\hat{\bar{\boldsymbol{t}}}_i^* = \left[\hat{\bar{t}}_{ig}^*\right]_{g=1}^G = \left[\frac{1}{N}\sum_{k\in s}\frac{1}{h_N}\mathcal{K}\left(\frac{x_j - x_i}{h_N}\right)\frac{z_{igk}^{\dagger}}{\pi_k^*}\right]_{g=1}^G = \left[\frac{1}{N}\sum_{k\in s}\frac{z_{igk}^*}{\pi_k^*}\right]_{g=1}^G.$$

Using a Taylor approximation, we have

$$\hat{m}_i^* = m_i + \frac{1}{N}\sum_{k\in U} z_{ik}\left(\frac{I_k}{\pi_k^*} - 1\right) + \left.\frac{\partial \hat{m}_i^*}{\partial \delta}\right|_{\hat{\bar{\boldsymbol{t}}}_i^* = \bar{\boldsymbol{t}}_i, \delta = 0} \frac{\delta}{N^2} + R_{iN}^*,$$

where R_{iN}^* is a remainder term of $\hat{m}_i^*$, and

$$z_{ik} = \sum_{g=1}^G \left.\frac{\partial \hat{m}_i^*}{\partial \hat{\bar{t}}_{ig}^*}\right|_{\hat{\bar{\boldsymbol{t}}}_i^* = \bar{\boldsymbol{t}}_i, \delta = 0} z_{igk}^*.$$

Lemma 8. *Under Conditions 2 and 5 - 10, the following results hold:*

(i) For $\kappa \ge 0$,

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{i\in U}\left(\frac{1}{2Nh_N}\sum_{j\in U} I_{\{x_i - h_N \le x_j \le x_i + h_N\}}\right)^{\kappa} < \infty.$$

(ii) There exists an $N^* > 0$, unrelated to x, such that $N \ge N^*$ implies

$$\sum_{j\in U} I_{\{|x - x_h| \le h_N\}} \ge q + 1.$$

(iii) $\{\bar{t}_{ig}\}$ are uniformly bounded in i and $\{\hat{\bar{t}}_{ig}^*\}$ are uniformly bounded in i and s.

(iv) $\{m_i\}$ are uniformly bounded in i and $\{\hat{m}_i^*\}$ are uniformly bounded in i and s.

(v) The first four order mixed partial derivatives of $\hat{m}_i^*$ with respect to $\bar{t}_{ig}$ and δ, evaluated at $\hat{\bar{\boldsymbol{t}}}_i^* = \bar{\boldsymbol{t}}_i, \delta = 0$, are uniformly bounded in i.

(vi) $\{R_{iN}^{*2}\}$ are uniformly bounded in i and s.

Proof. From Lemma 2 of Breidt and Opsomer (2000), the results (i) and (ii) are right. For (iii), we define $I_{i,k}(h_N) = I_{\{|x_k - x_i| \leq h_N\}}$, and so

$$\begin{aligned}\limsup_{N\to\infty} |\bar{t}_{ig}| &= \limsup_{N\to\infty} \left| \sum_{k\in U} \frac{1}{Nh_N} \mathcal{K}\left(\frac{x_k - x_i}{h_N}\right) (x_k - x_i)^{p_1} y_k^{p_2} \right| \\ &\leq \limsup_{N\to\infty} \frac{C}{Nh_N} \sum_{k\in U} I_{i,k}(h_N)\end{aligned}$$

and

$$\begin{aligned}\limsup_{N\to\infty} \left|\hat{\bar{t}}_{ig}^*\right| &= \limsup_{N\to\infty} \left| \sum_{k\in U} \frac{1}{Nh_N} \mathcal{K}\left(\frac{x_k - x_i}{h_N}\right) (x_k - x_i)^{p_1} y_k^{p_2} \frac{I_k}{\pi_k^*} \right| \\ &\leq \limsup_{N\to\infty} \frac{C}{N} \sum_{k\in s} \frac{I_{i,k}(h_N)}{h_N \pi_k}.\end{aligned}$$

Then using Lemma 1 of Breidt and Opsomer (2000) and Conditions 2 and 10, we see that $\{\bar{t}_{ig}\}$ are uniformly bounded in i, and $\{\hat{\bar{t}}_{ig}^*\}$ are uniformly bounded in i and s.

Similar to the proofs of Lemma 2 (iv)-(vi) of Breidt and Opsomer (2000), we can verify that the results (iv)-(vi) hold. This completes the proof of Lemma 8.

Lemma 9. *If Conditions 2 and 4 - 10 hold, then*

$$\frac{1}{N} \sum_{i\in U} E_p \left(\hat{\bar{t}}_{ig}^* - \bar{t}_{ig}\right)^4 = O\left(\frac{1}{n^2 h_N^2 \pi_{(K)}^4}\right).$$

Proof. By Condition 8, we observe that

$$\left|\mathcal{K}\left(\frac{x_k - x_i}{h_N}\right)(x_k - x_i)^{p_1} y_k^{p_2}\right| \leq C \cdot I_{i,k}(h_N).$$

Let $I_{i,kl}(h_N) = I_{i,k}(h_N) \cdot I_{i,l}(h_N)$, $I_{i,klp}(h_N) = I_{i,k}(h_N) \cdot I_{i,l}(h_N) \cdot I_{i,p}(h_N)$ and $I_{i,klpq}(h_N) = I_{i,k}(h_N) \cdot I_{i,l}(h_N) \cdot I_{i,p}(h_N) \cdot I_{i,q}(h_N)$. Then

$$\begin{aligned}
&\frac{1}{N}\sum_{i\in U} \mathrm{E}_p\left(\hat{\tilde{t}}^*_{ig} - \bar{t}_{ig}\right)^4 \\
&\quad \leq \frac{C}{N^5 h_N^4} \sum_{i,k,l,p,q\in U} \left| I_{i,klpq}(h_N) \cdot \mathrm{E}_p\left(\Phi_k^* \Phi_l^* \Phi_p^* \Phi_q^*\right) \right| \\
&\quad = \frac{C}{N^5 h_N^4} \sum_{i,k,l,p,q\in U} \left| I_{i,klpq}(h_N) \cdot \mathrm{E}_p\left[\left(\tilde{\Phi}_k + \Delta_k\right)\left(\tilde{\Phi}_l + \Delta_l\right)\left(\tilde{\Phi}_p + \Delta_p\right)\right.\right. \\
&\qquad \left.\left.\left(\tilde{\Phi}_q + \Delta_q\right)\right]\right| \\
&\quad \leq \frac{C}{N^5 h_N^4} \sum_{i,k,l,p,q\in U} \left\{ I_{i,klpq}(h_N) \cdot \left[\left|\mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p\tilde{\Phi}_q)\right| + \left|\mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p\Delta_q)\right|\right.\right. \\
&\qquad \left.\left. + \left|\mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\Delta_p\Delta_q)\right| + \left|\mathrm{E}_p(\tilde{\Phi}_k\Delta_l\Delta_p\Delta_q)\right| + \left|\Delta_k\Delta_l\Delta_p\Delta_q\right|\right]\right\} \\
&\quad = \frac{C}{N^5 h_N^4} \sum_{i\in U} \left\{ \left(\sum_{q\in U}\left|I_{i,q}(h_N)\cdot\Delta_q\right|\right)^4 + \sum_{k,l,p,q\in U}\left|I_{i,klpq}(h_N)\cdot \mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p\tilde{\Phi}_q)\right|\right. \\
&\qquad + \left(\sum_{q\in U}\left|I_{i,q}(h_N)\cdot\Delta_q\right|\right)\left(\sum_{k,l,p\in U}\left|I_{i,klp}(h_N)\cdot\mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p)\right|\right) \\
&\qquad \left. + \left(\sum_{q\in U}\left|I_{i,q}(h_N)\cdot\Delta_q\right|\right)^2\left(\sum_{k,l\in U}\left|I_{i,kl}(h_N)\cdot\mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l)\right|\right)\right\}. \qquad \text{(S1.20)}
\end{aligned}$$

By Definition 1, we have $K + 1 \leq \pi_{(K)}^{-1}$, and it follows that

$$\sum_{q\in U}\left|I_{i,q}(h_N)\cdot\Delta_q\right| = \sum_{q\in U_2}\left|\frac{\pi_q - \pi_q^*}{\pi_q^*}\right| = O\left(\pi_{(K)}^{-1}\right). \qquad \text{(S1.21)}$$

Following Lemma 3 of Breidt and Opsomer (2000), we have

$$\begin{aligned}
&\frac{1}{N^5h_N^4}\sum_{i,k,l,p,q\in U}\left|I_{i,klpq}(h_N)\cdot \mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p\tilde{\Phi}_q)\right| \\
&\le \frac{1}{N^5h_N^4\pi_{(K)}^4}\sum_{i,k,l,p,q\in U}\left|I_{i,klpq}(h_N)\cdot \mathrm{E}_p[(I_k-\pi_k)(I_l-\pi_l)(I_p-\pi_p)(I_q-\pi_q)]\right| \\
&= O\left(\frac{1}{n^2h_N^2\pi_{(K)}^4}\right). \qquad \text{(S1.22)}
\end{aligned}$$

From Lemma 8 (i) and Conditions 2, 4 and 9, we obtain

$$\begin{aligned}
&\frac{1}{N^3h_N^2}\sum_{i,k,l\in U}\left|I_{i,kl}(h_N)\cdot \mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l)\right| \\
&\le \frac{1}{N\pi_{(K)}^2}\sum_{i\in U}\left\{\max_{k\neq l\in U}|\mathrm{E}_p[(I_k-\pi_k)(I_l-\pi_l)]|\cdot\left(\sum_{k\in U}\frac{I_{i,k}(h_N)}{Nh_N}\right)^2 \right.\\
&\qquad \left.+\frac{1}{Nh_N}\sum_{k\in U}\frac{I_{i,k}(h_N)}{Nh_N}\right\} \\
&= O\left(\frac{1}{n\pi_{(K)}^2}+\frac{1}{Nh_N\pi_{(K)}^2}\right) \qquad \text{(S1.23)}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{N^4h_N^3}\sum_{i,k,l,p\in U}\left|I_{i,klp}(h_N)\cdot \mathrm{E}_p(\tilde{\Phi}_k\tilde{\Phi}_l\tilde{\Phi}_p)\right| \\
&\le \frac{1}{N\pi_{(K)}^3}\sum_{i\in U}\left\{\max_{k,l,p\in D_{3,N}}|\mathrm{E}_p[(I_k-\pi_k)(I_l-\pi_l)(I_p-\pi_p)]|\cdot\left(\sum_{k\in U}\frac{I_{i,k}(h_N)}{Nh_N}\right)^3\right. \\
&\quad \left.+\max_{k\neq l\in U}|\mathrm{E}_p[(I_k-\pi_k)(I_l-\pi_l)]|\cdot\frac{1}{Nh_N}\left(\sum_{k\in U}\frac{I_{i,k}(h_N)}{Nh_N}\right)^2+\frac{1}{N^2h_N^2}\left(\sum_{k\in U}\frac{I_{i,k}(h_N)}{Nh_N}\right)\right\} \\
&= O\left(\frac{1}{n\pi_{(K)}^3}\right). \qquad \text{(S1.24)}
\end{aligned}$$

Thus, combining Equations (S1.20) - (S1.24) and Condition 9, we see that

$$\frac{1}{N}\sum_{i\in U}\mathrm{E}_p\left(\hat{\hat{t}}^*_{ig}-\bar{t}_{ig}\right)^4=O\left(\frac{1}{n^2h_N^2\pi^4_{(K)}}\right).$$

This completes the proof of Lemma 9.

Lemma 10. *If Conditions 2 and 4 - 10 are satisfied, then*

$$\frac{1}{N}\sum_{i\in U}E_p\left(R^{*2}_{iN}\right)=O\left(\frac{1}{n^2h_N^2\pi^4_{(K)}}\right).$$

Proof. Using the Taylor linearization for the estimator $\hat{m}^*_i=f(\hat{\hat{\boldsymbol{t}}}^*_i,\delta)$ at the point $(\bar{\boldsymbol{t}}_i,0)$, we have

$$\begin{aligned}\hat{m}^*_i&=m_i+\sum_{g=1}^{G}\left.\frac{\partial\hat{m}^*_i}{\partial\hat{\hat{t}}^*_{ig}}\right|_{\hat{\hat{\boldsymbol{t}}}^*_i=\bar{\boldsymbol{t}}_i,\delta=0}(\hat{\hat{t}}^*_{ig}-\bar{t}_{ig})+\left.\frac{\partial\hat{m}^*_i}{\partial\delta}\right|_{\hat{\hat{\boldsymbol{t}}}^*_i=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2}+R^*_{iN}\\&=m_i+\frac{1}{N}\sum_{k\in U}z_{ik}\left(\frac{I_k}{\pi^*_k}-1\right)+\left.\frac{\partial\hat{m}^*_i}{\partial\delta}\right|_{\hat{\hat{\boldsymbol{t}}}^*_i=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2}+R^*_{iN}.\end{aligned}$$

Note that R^{*2}_{iN} and its first three derivatives at $(\bar{\boldsymbol{t}}_i,0)$ evaluate to zero. So similar to the proof of Lemma 3, by Lemma 9 and using Theorem 5.4.3 of Fuller (1996) with $\alpha=1,s=4$ and $a_N^4=O\left(n^{-2}h_N^{-2}\pi^{-4}_{(K)}\right)$, we obtain that

$$\frac{1}{N}\sum_{i\in U}\mathrm{E}_p\left(R^{*2}_{iN}\right)=O\left(\frac{1}{n^2h_N^2\pi^4_{(K)}}\right).$$

This completes the proof of Lemma 10.

Lemma 11. *If Conditions 2 and 4 - 10 hold, then*

$$E_p(\tilde{b}^2_N)\triangleq E_p\left\{\frac{1}{N}\sum_{k\in U}\frac{(\pi^*_k-I_k)(\hat{m}^*_k-m_k)}{\pi^*_k}\right\}^2=O\left(\frac{1}{n^2h_N^2\lambda_N^6}\right).$$

Proof. Using Lemma 8 (v) and Loève's c_r inequality, we have

$$\begin{aligned}
\mathrm{E}_p(\tilde{b}_N^2) &= \mathrm{E}_p\left\{\frac{1}{N}\sum_{i\in U}\left(1-\frac{I_i}{\pi_i^*}\right)\left(\frac{1}{N}\sum_{k\in U} z_{ik}\left(\frac{I_k}{\pi_k^*}-1\right)+\left.\frac{\partial \hat{m}_i^*}{\partial \delta}\right|_{\hat{\bar{\boldsymbol{t}}}_i^*=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2}+R_{iN}^*\right)\right\}^2 \\
&= \mathrm{E}_p\left\{\frac{1}{N^2}\sum_{i,k\in U}\frac{z_{ik}(\pi_i^*-I_i)(I_k-\pi_k^*)}{\pi_i^*\pi_k^*}+\frac{1}{N}\sum_{i\in U}\left(1-\frac{I_i}{\pi_i^*}\right)\left.\frac{\partial \hat{m}_i^*}{\partial \delta}\right|_{\hat{\bar{\boldsymbol{t}}}_i^*=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2}\right. \\
&\quad \left. +\frac{1}{N}\sum_{i\in U}\frac{(\pi_i^*-I_i)R_{iN}^*}{\pi_i^*}\right\}^2 \\
&\leq C\mathrm{E}_p\left\{\frac{1}{N^2}\sum_{i,k\in U}\frac{z_{ik}(\pi_i^*-I_i)(I_k-\pi_k^*)}{\pi_i^*\pi_k^*}\right\}^2+C\mathrm{E}_p\left\{\frac{1}{N}\sum_{i\in U}\frac{(\pi_i^*-I_i)R_{iN}^*}{\pi_i^*}\right\}^2 \\
&\quad +O\left(\frac{\delta^2}{N^4}\right) \\
&\triangleq \tilde{b}_{1N}+\tilde{b}_{2N}+O\left(\frac{\delta^2}{N^4}\right).
\end{aligned}$$

By Lemma 8 (v), it is clear that

$$\max_{i,k\in U} z_{ik}^2=\max_{i,k\in U}\left\{\sum_{g=1}^{G}\left.\frac{\partial \hat{m}_i^*}{\partial \hat{\bar{t}}_{ig}^*}\right|_{\hat{\bar{\boldsymbol{t}}}_i^*=\bar{\boldsymbol{t}}_i,\delta=0}\frac{1}{h_N}\mathcal{K}\left(\frac{x_j-x_i}{h_N}\right)z_{igk}^{\dagger}\right\}^2\leq Ch_N^{-2}.$$

Combining Lemma 1 and the fact that $\pi_{(K)} \geq \lambda_N$, we see that $\tilde{b}_{1N} = O\left(n^{-2}h_N^{-2}\lambda_N^{-4}\right)$. Further, using Lemma 10 and Cauchy-Schwarz inequality, we obtain

$$\tilde{b}_{2N}\leq\frac{C}{N}\mathrm{E}_p\left\{\sum_{i\in U}\frac{R_{iN}^{*2}(\pi_i^*-I_i)^2}{\pi_i^{*2}}\right\}\leq\frac{C}{N\lambda_N^2}\sum_{i\in U}\mathrm{E}_p\big(R_{iN}^{*2}\big)=O\left(n^{-2}h_N^{-2}\lambda_N^{-6}\right).$$

Hence $E(\tilde{b}_N^2)=O\left(n^{-2}h_N^{-2}\lambda_N^{-6}\right)$. This completes the proof of Lemma 11.

Lemma 12. *If Conditions 2 and 4 - 10 hold, and* $(nh_N\lambda_N^2)^{-1}=o(1)$, *then*

$$\frac{1}{N}\sum_{i\in U}E_p\left(\hat{m}_i^*-m_i\right)^2=O\left(\frac{1}{nh_N\lambda_N^2}\right)$$

and

$$\frac{1}{K}\sum_{i\in U_2} E_p\left(\hat{m}_i^* - m_i\right)^2 = O\left(\frac{1}{nh_N\lambda_N^2}\right).$$

Proof. From Lemma 10, it can be seen that

$$\begin{aligned}
\mathrm{E}_p\left(\hat{m}_i^* - m_i\right)^2 &= \mathrm{E}_p\left\{\frac{1}{N}\sum_{k\in U} z_{ik}\left(\frac{I_k}{\pi_k^*}-1\right) + \left.\frac{\partial \hat{m}_i^*}{\partial\delta}\right|_{\hat{\boldsymbol{t}}_i^*=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2} + R_{iN}^*\right\}^2 \\
&\le C\mathrm{E}_p\left\{\frac{1}{N}\sum_{k\in U} z_{ik}\left(\frac{I_k}{\pi_k^*}-1\right)\right\}^2 + C\mathrm{E}_p\left\{\left.\frac{\partial \hat{m}_i^*}{\partial\delta}\right|_{\hat{\boldsymbol{t}}_i^*=\bar{\boldsymbol{t}}_i,\delta=0}\frac{\delta}{N^2}\right\}^2 \\
&\quad + C\mathrm{E}_p\left(R_{iN}^{*2}\right). \qquad \text{(S1.25)}
\end{aligned}$$

Using Lemma 8 (v), the second term of (S1.25) is $O(\delta N^{-2})$. By Lemma 10 and the fact that $\lambda_N \le \pi_{(K)}$, we obtain

$$\frac{1}{N}\sum_{i\in U}\mathrm{E}_p\left(R_{iN}^{*2}\right) = O\left(\frac{1}{n^2h_N^2\lambda_N^4}\right). \qquad \text{(S1.26)}$$

From Lemma 8 (i), we have

$$\frac{1}{N^2}\sum_{i,k\in U} z_{ik}^2 \le \frac{C}{Nh_N}\sum_{i\in U}\frac{1}{Nh_N}\sum_{k\in U} I_{i,k}(h_N) = O\left(\frac{1}{h_N}\right).$$

Thus, using Condition 9 and $K \le \lambda_N^{-1}$, we see that

$$\begin{aligned}
&\frac{1}{N}\sum_{i\in U}\mathrm{E}_p\left\{\frac{1}{N}\sum_{k\in U} z_{ik}\left(\frac{I_k}{\pi_k^*}-1\right)\right\}^2 \\
&\le \frac{1}{N}\sum_{i\in U}\sum_{k,l\in U}\left(\frac{\pi_{kl}-\pi_k\pi_l^*-\pi_k^*\pi_l+\pi_k^*\pi_l^*}{N^2\pi_{(K)}^2}z_{ik}z_{il}\right) \\
&\le \frac{C}{N^3\pi_{(K)}^2}\sum_{i\in U}\left\{\sum_{k\in U} z_{ik}^2 + \frac{1}{n}\sum_{k\neq l\in U_1}|z_{ik}z_{il}| + \sum_{k\neq l\in U_2}|z_{ik}z_{il}| + \frac{1}{n}\sum_{k\in U_1;l\in U_2}|z_{ik}z_{il}|\right\}
\end{aligned}$$

$$\leq \frac{C}{N^3\pi_{(K)}^2}\sum_{i\in U}\left\{\left(1+\frac{N}{n}\right)\sum_{k\in U} z_{ik}^2 + \frac{K^2}{h_N^2} + \frac{NK}{nh_N^2}\right\}$$
$$= O\left(\frac{1}{nh_N\lambda_N^2}\right). \tag{S1.27}$$

Combining (S1.25) - (S1.27) and $(nh_N\lambda_N^2)^{-1} = o(1)$, we obtain

$$\frac{1}{N}\sum_{i\in U}\mathrm{E}_p\left(\hat{m}_i^* - m_i\right)^2 = O\left(\frac{1}{nh_N\lambda_N^2}\right).$$

Similarly, we can prove that

$$\frac{1}{K}\sum_{i\in U_2}\mathrm{E}_p\left(\hat{m}_i^* - m_i\right)^2 = O\left(\frac{1}{nh_N\lambda_N^2}\right).$$

This completes the proof of Lemma 12.

S1.10 Proof of Theorem 8

We first show that the improved estimator $\tilde{t}_{yr}^*$ is asymptotically design unbiased and design consistent. By Markov's inequality, it suffices to prove that $\mathrm{E}_p\left|\tilde{t}_{yr}^* - \bar{t}_y\right| = o(1)$. Note that

$$\tilde{t}_{yr}^* - \bar{t}_y = \frac{1}{N}\sum_{k\in U}\frac{(I_k - \pi_k^*)(y_k - m_k)}{\pi_k^*} + \frac{1}{N}\sum_{k\in U}\frac{(\pi_k^* - I_k)(\hat{m}_k^* - m_k)}{\pi_k^*}$$
$$\triangleq \tilde{a}_N + \tilde{b}_N.$$

Using Lemmas 2 and 8 (iv), we have

$$\mathrm{E}_p\left(\tilde{a}_N^2\right) = \mathrm{E}_p\left\{\frac{1}{N}\sum_{k\in U}\frac{(I_k - \pi_k^*)(y_k - m_k)}{\pi_k^*}\right\}^2$$
$$= \frac{1}{N^2}\sum_{k,l\in U}\left\{\frac{\pi_{kl} - \pi_k\pi_l^* - \pi_k^*\pi_l + \pi_k^*\pi_l^*}{\pi_k^*\pi_l^*}(y_k - m_k)(y_l - m_l)\right\}$$

$$= O\left(n^{-1}\lambda_N^{-2}\right). \tag{S1.28}$$

Further, from Lemma 11, we see that $\mathrm{E}_p\left(\tilde{b}_N^2\right) = O\left(n^{-2}h_N^{-2}\lambda_N^{-6}\right)$. Thus, using Cauchy-Schwarz inequality and the condition $(nh_N^2\lambda_N^4)^{-1} = o(1)$, we obtain

$$\mathrm{E}_p\left|\tilde{t}_{yr}^* - \bar{t}_y\right| \le \mathrm{E}_p\left|\tilde{a}_N\right| + \mathrm{E}_p\left|\tilde{b}_N\right| = O(n^{-1/2}\lambda_N^{-1}).$$

Next, we turn to evaluate the design mean squared error of $\tilde{t}_{yr}^*$. Observe that

$$\mathrm{MSE}_p\left(\tilde{t}_{yr}^*\right) = \mathrm{E}_p\left(\tilde{t}_{yr}^* - \bar{t}_y\right)^2 = \mathrm{E}_p(\tilde{a}_N^2) + \mathrm{E}_p(\tilde{b}_N^2) + 2\mathrm{E}_p(\tilde{a}_N\tilde{b}_N).$$

Combining (S1.28), Lemma 11 and Cauchy-Schwarz inequality, we have

$$\mathrm{E}_p|\tilde{a}_N\tilde{b}_N| \le \left\{\mathrm{E}_p(\tilde{a}_N^2)\mathrm{E}_p(\tilde{b}_N^2)\right\}^{1/2} = O(n^{-3/2}h_N^{-1}\lambda_N^{-4}).$$

Hence, by the condition $(n\lambda_N^4h_N^2)^{-1} = o(1)$, the design mean squared error of $\tilde{t}_{yr}^*$ is

$$\begin{aligned}\mathrm{MSE}_p\left(\tilde{t}_{yr}^*\right) &= \mathrm{E}_p(\tilde{a}_N^2) + O(n^{-3/2}h_N^{-1}\lambda_N^{-4}) \qquad \text{(S1.29)}\\ &= \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}(y_k - m_k)(y_l - m_l) + O(n^{-3/2}h_N^{-1}\lambda_N^{-4}).\end{aligned}$$

Finally, to prove that $\widehat{\mathrm{AMSE}}_p\left(\tilde{t}_{yr}^*\right)$ is an asymptotically design unbiased estimator, it suffices to show that

$$\mathrm{E}_p\left|\widehat{\mathrm{AMSE}}_p\left(\tilde{t}_{yr}^*\right) - \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}(y_k - m_k)(y_l - m_l)\right| = O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right).$$

Define

$$\tilde{A}_N = \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ (y_k - m_k)(y_l - m_l) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l - \pi_{kl}}{\pi_{kl}} \right\} \right|$$

and

$$\tilde{B}_N = \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ 2\left(y_k - m_k\right)\left(m_l - \hat{m}_l^*\right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l}{\pi_{kl}} \right\} \right| + \mathrm{E}_p \left| \frac{1}{N^2} \sum_{k,l \in U} \left\{ \left(m_k - \hat{m}_k^*\right)\left(m_l - \hat{m}_l^*\right) \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} \frac{I_k I_l}{\pi_{kl}} \right\} \right|.$$

Then

$$\mathrm{E}_p \left| \widehat{\mathrm{AMSE}}(\tilde{t}_{yr}^*) - \frac{1}{N^2} \sum_{k,l \in U} \frac{\Delta_{kl}^*}{\pi_k^* \pi_l^*} (y_k - m_k)(y_l - m_l) \right| \leq \tilde{A}_N + \tilde{B}_N.$$

By Condition 2 and Lemma 8 (iv), we see that

$$\max_{k,l \in U} \left| \frac{\left(y_k - m_k\right)\left(y_l - m_l\right)}{\pi_k^* \pi_l^* \pi_{kl}} \right| \leq \frac{C}{\lambda_N^2 \lambda_N^*}.$$

Combining Definition 1 and the condition $(nh_N^2 \lambda_N^8)^{-1} = o(1)$, we have $\sqrt{n}K/N = O(1)$. Thus, $\tilde{A}_N = O\left(n^{-1-\kappa} \lambda_N^{-2} \lambda_N^{*-1}\right)$ is obtained from Lemma 5, where $\kappa = \min\left\{\frac{\alpha}{2}, \frac{1}{4}\right\}$.

Now we prove $\tilde{B}_N = O\left(n^{-2} \lambda_N^{-4} \lambda_N^{*-1} h_N^{-1}\right)$. By Lemma 8 (iv), it is seen that

$$\max_{k,l \in U} \left| \frac{2 I_k I_l}{\pi_k^* \pi_l^* \pi_{kl}} \left(y_k - m_k\right) \right| \leq \frac{C}{\lambda_N^2 \lambda_N^*}.$$

Then, using Lemmas 6 and 12 and the fact of $\sqrt{n}K/N = O(1)$, the order of the first term in $\tilde{B}_N$ is $O\left(n^{-2} \lambda_N^{-4} \lambda_N^{*-1} h_N^{-1}\right)$. Similarly, it is straightforward to show that the orders of the other terms in $\tilde{B}_N$ are $O\left(n^{-2} \lambda_N^{-4} \lambda_N^{*-1} h_N^{-1}\right)$

too.

Thus, by the condition $(nh_N^2\lambda_N^8)^{-1} = o(1)$, we obtain

$$\mathrm{E}_p \left| \widehat{\mathrm{AMSE}}_p\left(\tilde{t}_{yr}^*\right) - \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}(y_k - m_k)(y_l - m_l)\right| = O\left(n^{-1-\kappa}\lambda_N^{-2}\lambda_N^{*-1}\right).$$

This completes the proof of Theorem 8.

S1.11 Proof of Theorem 9

From Breidt and Opsomer (2000), the design mean squared error of the estimator $\tilde{t}_{yr}$ is

$$\mathrm{MSE}_p(\tilde{t}_{yr}) = \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}}{\pi_k\pi_l}(y_k - m_k)(y_l - m_l) + o\left(n^{-1}\right), \tag{S1.30}$$

where $\Delta_{kl} = \pi_{kl} - \pi_k\pi_l$. By Theorem 2, we see that

$$\begin{aligned}
&\frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}^*}{\pi_k^*\pi_l^*}(y_k - m_k)(y_l - m_l)\\
&\quad = E_p\left\{\frac{1}{N}\sum_{k\in s}\frac{y_k - m_k}{\pi_k^*} - \frac{1}{N}\sum_{k\in U}(y_k - m_k)\right\}^2\\
&\quad \leq E_p\left\{\frac{1}{N}\sum_{k\in s}\frac{y_k - m_k}{\pi_k} - \frac{1}{N}\sum_{k\in U}(y_k - m_k)\right\}^2 + o\left(n^{-1}\right)\\
&\quad = \frac{1}{N^2}\sum_{k,l\in U}\frac{\Delta_{kl}}{\pi_k\pi_l}(y_k - m_k)(y_l - m_l) + o\left(n^{-1}\right).
\end{aligned}$$

Thus, combining (S1.29), (S1.30) and the condition $(n\lambda_N^8 h_N^2)^{-1} = o(1)$, we obtain

$$\mathrm{MSE}_p\left(\tilde{t}_{yr}^*\right) \leq \mathrm{MSE}_p\left(\tilde{t}_{yr}\right) + o\left(n^{-1}\right).$$

This completes the proof of Theorem 9.

S2 Additional numerical results

S2.1 Bias and variance

In this subsection, we report the biases and variances of all estimators under various scenarios.

For Section 7.1, Tables S1 - S3 show the squared biases corresponding to Tables 1 - 3, and Tables S4 - S6 show the variances corresponding to Tables 1 - 3;

For Section 7.2, Tables S7 - S9 show the squared biases corresponding to Tables 4 - 6, and Tables S10 - S12 show the variances corresponding to Tables 4 - 6.

For Section 7.3, Tables S13 - S14 show the squared biases and variances corresponding to Table 7 respectively.

Table S1: Empirical squared bias of each estimator under different f and $\boldsymbol{\rho}$.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	1.00E-03	6.63E-04	1.07E-03	1.16E-03	**3.43E-06**	4.15E-05
HT-Beta	1.20E-01	1.72E-01	1.99E-01	2.18E-01	2.25E-01	2.24E-01
HT-MSE	**3.22E-04**	**3.84E-04**	**6.43E-04**	**8.01E-04**	1.12E-04	**4.05E-06**
HT-Rob	1.49E-01	9.83E-02	7.84E-02	6.70E-02	4.68E-02	4.05E-02
IHT	2.12E-01	1.04E-01	7.28E-02	5.36E-02	4.43E-02	3.30E-02
HA	**6.67E-05**	**8.20E-05**	**5.30E-05**	**2.31E-05**	**9.70E-06**	**4.72E-06**
HA-Beta	1.35E-04	1.50E-04	1.47E-04	9.65E-05	1.30E-04	1.20E-04
HA-MSE	3.36E-04	2.09E-04	1.70E-04	9.35E-05	1.01E-04	7.71E-05
HA-Rob	2.08E-04	2.25E-04	1.53E-04	9.44E-05	7.26E-05	5.96E-05
IHA	1.53E-04	9.03E-05	5.86E-05	4.74E-05	5.32E-05	4.71E-05
LMA	2.94E-08	1.26E-05	1.25E-05	5.15E-07	1.49E-08	**5.89E-07**
LMA-Trim	6.27E-03	2.17E-03	1.55E-03	8.03E-04	8.81E-04	6.45E-04
LMA-Rob	**8.21E-10**	1.30E-05	8.86E-06	**1.93E-07**	2.67E-09	1.49E-06
ILMA	7.17E-08	**4.58E-06**	**4.94E-06**	2.19E-07	**9.25E-10**	6.86E-07
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	**5.60E-03**	**2.90E-06**	**6.40E-04**	**2.86E-04**	**1.47E-04**	**2.48E-04**
HT-Beta	4.80E-02	8.78E-02	1.08E-01	1.05E-01	1.18E-01	1.13E-01
HT-MSE	3.41E-02	7.25E-03	2.18E-03	5.91E-03	1.83E-03	4.15E-03
HT-Rob	4.15E-02	3.98E-02	3.69E-02	2.19E-02	2.42E-02	1.52E-02
IHT	8.86E-02	4.73E-02	3.40E-02	2.16E-02	2.07E-02	1.36E-02
HA	**1.53E-03**	**2.77E-04**	**4.25E-04**	**2.10E-04**	**2.58E-04**	**1.06E-04**
HA-Beta	4.80E-03	4.26E-03	6.11E-03	5.69E-03	6.44E-03	5.91E-03
HA-MSE	1.56E-02	7.77E-03	6.94E-03	4.69E-03	4.31E-03	2.99E-03
HA-Rob	7.18E-03	3.31E-03	3.12E-03	2.13E-03	2.06E-03	1.40E-03
IHA	5.44E-03	1.88E-03	1.84E-03	1.14E-03	1.24E-03	7.76E-04
LMA	1.99E-06	3.85E-05	1.60E-05	**1.55E-09**	**2.29E-10**	3.87E-06
LMA-Trim	5.22E-03	3.69E-03	4.56E-03	4.84E-03	5.11E-03	5.08E-03
LMA-Rob	**8.80E-07**	**1.81E-05**	**1.01E-05**	1.97E-07	6.58E-07	**1.49E-06**
ILMA	2.60E-06	4.67E-05	1.53E-05	2.30E-06	9.51E-08	6.79E-06
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	**2.12E-04**	**1.49E-04**	**8.57E-06**	**8.53E-07**	**1.37E-06**	**1.14E-04**
HT-Beta	1.24E-02	2.08E-02	2.35E-02	2.54E-02	2.46E-02	2.72E-02
HT-MSE	1.95E-02	9.70E-03	8.74E-03	7.46E-03	5.44E-03	3.08E-03
HT-Rob	2.00E-02	1.28E-02	8.35E-03	6.39E-03	5.07E-03	5.35E-03
IHT	2.89E-02	1.22E-02	5.96E-03	4.88E-03	3.63E-03	3.35E-03
HA	**4.70E-03**	**2.27E-03**	**5.61E-04**	**4.45E-04**	**2.11E-04**	**3.38E-04**
HA-Beta	1.25E-02	1.29E-02	1.07E-02	1.17E-02	1.03E-02	1.13E-02
HA-MSE	3.93E-02	2.16E-02	1.26E-02	1.03E-02	7.16E-03	6.63E-03
HA-Rob	1.95E-02	1.15E-02	6.11E-03	5.04E-03	3.47E-03	3.49E-03
IHA	1.53E-02	7.87E-03	3.58E-03	3.21E-03	2.04E-03	1.92E-03
LMA	**3.01E-09**	6.06E-06	1.03E-06	2.77E-07	3.85E-07	1.42E-07
LMA-Trim	3.77E-03	4.89E-03	5.47E-03	7.06E-03	6.80E-03	7.19E-03
LMA-Rob	8.43E-07	**2.23E-06**	2.77E-06	**2.04E-07**	1.53E-06	1.88E-06
ILMA	1.57E-07	4.33E-06	**4.24E-07**	1.89E-06	**3.13E-08**	**3.74E-09**

[1] For each estimation type, the minimum Bias^2 of all estimators is displayed in bold.

Table S2: Empirical squared bias of each estimator in the case of 50 classes.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	**4.70E-04**	**2.39E-05**	**1.04E-06**	7.81E-04	8.62E-05	**4.00E-05**
HT-Beta	6.07E-02	9.17E-02	1.28E-01	1.40E-01	1.41E-01	1.36E-01
HT-MSE	1.31E-03	2.33E-04	8.97E-05	**4.60E-04**	**7.27E-06**	1.40E-04
HT-Rob	1.01E-01	6.27E-02	5.13E-02	4.80E-02	3.43E-02	2.46E-02
IHT	1.73E-01	4.54E-02	3.21E-02	3.50E-02	1.29E-02	1.02E-02
HA	**5.67E-05**	**2.13E-05**	**4.42E-05**	**2.08E-05**	**1.18E-06**	**6.87E-06**
HA-Beta	1.09E-04	9.02E-05	2.00E-04	1.12E-04	7.16E-05	9.14E-05
HA-MSE	4.84E-04	2.03E-04	2.71E-04	1.19E-04	5.85E-05	6.31E-05
HA-Rob	2.84E-04	1.52E-04	1.96E-04	8.87E-05	3.85E-05	4.42E-05
IHA	1.34E-04	5.32E-05	1.16E-04	5.46E-05	7.34E-06	1.74E-05
LMA	**8.23E-05**	**6.85E-06**	**3.97E-05**	**2.83E-05**	**9.86E-06**	**1.17E-06**
LMA-Trim	4.49E-03	1.46E-03	9.12E-04	6.82E-04	4.50E-04	4.29E-04
LMA-Rob	9.25E-05	1.86E-05	6.40E-05	5.36E-05	2.82E-05	1.01E-05
ILMA	1.02E-04	1.41E-05	4.91E-05	3.80E-05	1.31E-05	3.41E-06
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	**4.37E-04**	**3.07E-04**	**1.72E-05**	**6.88E-07**	**1.15E-04**	**1.78E-04**
HT-Beta	3.01E-02	5.78E-02	7.11E-02	8.16E-02	7.99E-02	7.65E-02
HT-MSE	1.79E-02	1.05E-02	4.99E-03	3.49E-03	1.50E-03	3.41E-03
HT-Rob	4.60E-02	3.03E-02	2.46E-02	2.05E-02	1.86E-02	1.11E-02
IHT	6.44E-02	3.69E-02	2.08E-02	1.05E-02	1.09E-02	7.29E-03
HA	**1.25E-03**	**5.29E-04**	**2.17E-04**	**6.95E-05**	**1.25E-04**	**1.32E-04**
HA-Beta	4.48E-03	5.17E-03	4.47E-03	4.24E-03	4.33E-03	4.82E-03
HA-MSE	1.54E-02	9.03E-03	5.53E-03	3.81E-03	3.26E-03	2.93E-03
HA-Rob	6.82E-03	4.15E-03	2.57E-03	1.71E-03	1.65E-03	1.54E-03
IHA	5.01E-03	3.24E-03	1.58E-03	7.47E-04	9.10E-04	9.42E-04
LMA	**1.18E-05**	**4.57E-05**	**7.95E-07**	1.27E-06	**2.34E-07**	**1.03E-05**
LMA-Trim	3.01E-03	3.28E-03	2.71E-03	2.85E-03	3.00E-03	3.90E-03
LMA-Rob	1.39E-05	5.78E-05	2.80E-06	**7.71E-11**	2.62E-06	1.43E-05
ILMA	2.64E-05	6.20E-05	3.73E-06	2.41E-09	3.37E-06	1.72E-05
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	**1.67E-04**	**9.78E-05**	**5.06E-05**	**1.73E-04**	**1.05E-06**	**1.77E-05**
HT-Beta	8.60E-03	1.42E-02	1.56E-02	1.51E-02	1.72E-02	1.87E-02
HT-MSE	1.91E-02	1.06E-02	6.72E-03	8.82E-03	4.48E-03	3.80E-03
HT-Rob	1.63E-02	9.38E-03	6.21E-03	2.90E-03	3.42E-03	3.08E-03
IHT	2.20E-02	5.44E-03	2.63E-03	1.71E-03	9.76E-04	1.12E-03
HA	**4.08E-03**	**1.29E-03**	**7.13E-04**	**2.91E-04**	**3.98E-04**	**1.39E-04**
HA-Beta	1.17E-02	1.10E-02	1.01E-02	9.84E-03	9.80E-03	9.57E-03
HA-MSE	3.67E-02	1.92E-02	1.21E-02	8.66E-03	7.24E-03	5.63E-03
HA-Rob	1.87E-02	9.84E-03	6.45E-03	4.44E-03	3.93E-03	2.85E-03
IHA	1.45E-02	5.37E-03	3.13E-03	2.64E-03	1.68E-03	1.26E-03
LMA	**4.51E-05**	**1.05E-05**	**2.39E-05**	**2.76E-05**	**2.00E-05**	**2.90E-06**
LMA-Trim	4.13E-03	5.55E-03	6.11E-03	7.32E-03	7.20E-03	7.23E-03
LMA-Rob	5.29E-05	1.55E-05	3.25E-05	3.68E-05	2.88E-05	1.06E-05
ILMA	5.32E-05	2.36E-05	4.85E-05	5.29E-05	3.72E-05	1.82E-05

[1] For each estimation type, the minimum Bias^2 of all estimators is displayed in bold.

Table S3: Empirical squared bias of each estimator in the case of 100 classes.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	**1.14E-05**	**4.27E-05**	**3.07E-06**	**3.59E-05**	**1.05E-06**	**1.14E-03**
HT-Beta	9.72E-02	1.51E-01	1.72E-01	2.00E-01	1.97E-01	2.06E-01
HT-MSE	5.18E-04	4.49E-05	1.17E-04	2.19E-04	8.31E-05	1.59E-03
HT-Rob	1.28E-01	8.51E-02	6.17E-02	5.07E-02	4.18E-02	2.83E-02
IHT	1.89E-01	8.13E-02	5.82E-02	4.87E-02	3.22E-02	2.98E-02
HA	**1.03E-07**	**1.00E-04**	**1.25E-05**	**1.22E-07**	1.42E-06	2.35E-08
HA-Beta	2.45E-05	2.96E-04	1.82E-04	8.45E-05	7.96E-05	1.01E-04
HA-MSE	2.94E-04	5.41E-04	2.59E-04	8.30E-05	4.33E-05	3.88E-05
HA-Rob	9.81E-05	3.07E-04	1.08E-04	2.76E-05	1.30E-05	1.45E-05
IHA	3.42E-05	1.68E-04	5.02E-05	2.65E-06	**2.82E-07**	**3.47E-09**
LMA	**3.30E-05**	3.81E-06	7.16E-07	**2.49E-07**	4.11E-07	6.83E-07
LMA-Trim	3.75E-03	2.35E-03	1.42E-03	7.06E-04	4.80E-04	4.65E-04
LMA-Rob	3.43E-05	**1.18E-06**	**7.30E-09**	2.91E-06	**1.62E-07**	**1.18E-08**
ILMA	4.71E-05	2.33E-06	1.55E-06	2.91E-07	2.57E-07	2.84E-07
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	**6.61E-05**	**3.31E-05**	**5.11E-04**	**5.74E-05**	**4.82E-04**	**7.34E-06**
HT-Beta	4.23E-02	6.98E-02	8.78E-02	9.46E-02	9.29E-02	9.80E-02
HT-MSE	9.60E-03	7.61E-03	1.37E-03	3.63E-03	4.57E-03	2.23E-03
HT-Rob	5.88E-02	3.61E-02	3.38E-02	2.18E-02	1.58E-02	1.62E-02
IHT	8.06E-02	4.03E-02	2.86E-02	2.37E-02	1.31E-02	1.42E-02
HA	**1.74E-03**	**2.04E-04**	**3.96E-04**	**2.78E-04**	**8.92E-05**	**2.19E-06**
HA-Beta	4.89E-03	3.52E-03	4.69E-03	4.94E-03	4.02E-03	3.66E-03
HA-MSE	1.69E-02	7.11E-03	5.65E-03	4.38E-03	2.74E-03	1.86E-03
HA-Rob	7.70E-03	3.03E-03	3.03E-03	2.33E-03	1.40E-03	8.36E-04
IHA	5.30E-03	1.77E-03	1.52E-03	1.29E-03	5.76E-04	3.64E-04
LMA	1.41E-06	4.43E-05	**1.60E-05**	1.39E-06	**2.97E-05**	6.27E-05
LMA-Trim	3.33E-03	2.05E-03	2.61E-03	3.08E-03	2.68E-03	2.55E-03
LMA-Rob	**8.08E-09**	6.18E-05	2.45E-05	3.28E-08	4.45E-05	7.07E-05
ILMA	1.76E-07	**3.89E-05**	1.71E-05	**1.14E-08**	3.40E-05	**5.39E-05**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	**3.83E-04**	**1.33E-04**	**2.92E-05**	**1.05E-06**	**1.44E-04**	**2.89E-06**
HT-Beta	1.51E-02	2.26E-02	2.96E-02	3.20E-02	3.33E-02	3.20E-02
HT-MSE	4.53E-02	1.94E-02	1.02E-02	8.95E-03	4.73E-03	5.81E-03
HT-Rob	1.59E-02	1.02E-02	9.71E-03	7.48E-03	7.21E-03	4.66E-03
IHT	2.95E-02	1.11E-02	7.26E-03	6.06E-03	5.04E-03	2.45E-03
HA	**3.56E-03**	**1.20E-03**	**9.84E-04**	**3.62E-04**	**7.05E-04**	**1.84E-04**
HA-Beta	1.37E-02	1.23E-02	1.46E-02	1.34E-02	1.46E-02	1.30E-02
HA-MSE	4.12E-02	2.05E-02	1.60E-02	1.11E-02	1.02E-02	6.84E-03
HA-Rob	1.87E-02	9.67E-03	7.82E-03	5.10E-03	5.30E-03	3.32E-03
IHA	1.58E-02	6.58E-03	4.60E-03	2.97E-03	2.94E-03	1.40E-03
LMA	6.61E-07	3.29E-05	2.38E-07	3.44E-06	6.01E-06	**1.48E-06**
LMA-Trim	5.07E-03	5.50E-03	7.73E-03	7.72E-03	8.53E-03	8.73E-03
LMA-Rob	**5.16E-07**	**2.42E-05**	**1.35E-08**	4.21E-06	**5.62E-06**	2.06E-06
ILMA	1.38E-06	3.25E-05	2.65E-06	**2.49E-06**	1.03E-05	2.85E-06

[1] For each estimation type, the minimum Bias^2 of all estimators is displayed in bold.

Table S4: Empirical variance of each estimator under different f and $\boldsymbol{\rho}$.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	3.55192	1.39325	0.89120	0.61536	1.26330	0.53865
HT-Beta	1.00103	0.42025	**0.28565**	**0.20099**	**0.15146**	**0.11905**
HT-MSE	3.59888	1.38941	0.92091	0.62155	1.28192	0.56855
HT-Rob	1.58003	0.69933	0.48879	0.35023	0.50619	0.28948
IHT	**0.54432**	**0.36051**	0.29576	0.24146	0.20729	0.17411
HA	0.07496	0.04172	0.02975	0.02292	0.01916	0.01630
HA-Beta	0.06362	0.03060	0.02145	**0.01577**	**0.01140**	**0.00965**
HA-MSE	**0.05338**	**0.02848**	**0.02095**	0.01610	0.01222	0.01067
HA-Rob	0.06386	0.03433	0.02447	0.01881	0.01484	0.01274
IHA	0.06115	0.03357	0.02432	0.01890	0.01449	0.01247
LMA	0.02004	0.01032	0.00705	0.00524	0.00411	0.00339
LMA-Trim	0.07847	0.03325	0.02362	0.01686	0.01217	0.01032
LMA-Rob	0.01959	0.00979	**0.00662**	**0.00487**	**0.00375**	**0.00307**
ILMA	**0.01949**	**0.00966**	0.00662	0.00490	0.00383	0.00312
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	2.74296	0.73299	0.47299	0.50717	0.28961	0.25849
HT-Beta	0.46982	0.19709	**0.12755**	**0.09239**	**0.06492**	**0.05547**
HT-MSE	3.17825	0.91040	0.62006	0.57860	0.35553	0.30753
HT-Rob	1.01600	0.35634	0.23789	0.23334	0.15351	0.13681
IHT	**0.24816**	**0.17051**	0.14037	0.11537	0.09314	0.08224
HA	0.09368	0.05315	0.04047	0.03293	0.02606	0.02086
HA-Beta	0.07638	0.03821	**0.02590**	**0.02021**	**0.01551**	**0.01208**
HA-MSE	**0.06463**	**0.03581**	0.02602	0.02138	0.01670	0.01347
HA-Rob	0.07414	0.04102	0.03022	0.02470	0.01929	0.01545
IHA	0.07012	0.03882	0.02812	0.02246	0.01750	0.01378
LMA	0.04854	0.02294	0.01615	0.01154	0.00979	0.00762
LMA-Trim	0.08583	0.03739	0.02504	0.01825	0.01508	0.01121
LMA-Rob	**0.04748**	**0.02191**	0.01531	0.01061	0.00920	0.00704
ILMA	0.04767	0.02201	**0.01515**	**0.01037**	**0.00906**	**0.00683**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	0.28311	0.14866	0.09677	0.08055	0.06263	0.04678
HT-Beta	0.17257	0.07490	**0.04609**	**0.03303**	**0.02649**	**0.02016**
HT-MSE	0.53735	0.24907	0.15380	0.12735	0.09491	0.06737
HT-Rob	0.16366	0.09072	0.06232	0.05091	0.04022	0.03250
IHT	**0.09041**	**0.06085**	0.04680	0.03664	0.03054	0.02616
HA	0.10764	0.05936	0.04323	0.03452	0.02745	0.02331
HA-Beta	0.08608	0.04115	**0.02832**	**0.02016**	**0.01599**	**0.01324**
HA-MSE	0.07028	0.03996	0.02880	0.02141	0.01809	0.01577
HA-Rob	0.08046	0.04405	0.03162	0.02469	0.02022	0.01742
IHA	**0.06994**	**0.03841**	0.02861	0.02142	0.01791	0.01590
LMA	0.02501	0.01180	0.00817	0.00568	0.00465	0.00389
LMA-Trim	0.04537	0.02099	0.01523	0.01114	0.00934	0.00764
LMA-Rob	0.02477	0.01145	**0.00776**	**0.00532**	**0.00434**	**0.00361**
ILMA	**0.02431**	**0.01126**	0.00781	0.00534	0.00442	0.00367

[1] For each estimation type, the minimum variance of all estimators is displayed in bold.

Table S5: Empirical variance of each estimator in the case of 50 classes.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	2.20015	0.87810	0.64656	0.44185	0.36769	0.29738
HT-Beta	1.08996	0.43814	**0.29148**	**0.19201**	**0.15225**	**0.12279**
HT-MSE	2.14460	0.87363	0.64432	0.43505	0.36040	0.29433
HT-Rob	1.20527	0.54203	0.41732	0.29583	0.25338	0.20963
IHT	**0.51255**	**0.38474**	0.31847	0.21440	0.22630	0.17730
HA	0.08769	0.04610	0.03325	0.02313	0.02082	0.01612
HA-Beta	0.07189	0.03544	0.02295	**0.01539**	**0.01375**	**0.01024**
HA-MSE	**0.06025**	**0.03304**	**0.02226**	0.01561	0.01435	0.01092
HA-Rob	0.07347	0.03869	0.02685	0.01885	0.01707	0.01295
IHA	0.06703	0.03784	0.02613	0.01790	0.01712	0.01280
LMA	0.02211	0.01105	0.00769	0.00531	0.00445	0.00377
LMA-Trim	0.08110	0.03561	0.02257	0.01484	0.01311	0.00975
LMA-Rob	0.02167	**0.01065**	**0.00731**	0.00497	**0.00415**	0.00346
ILMA	**0.02155**	0.01069	0.00734	**0.00493**	0.00417	**0.00344**
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	1.11052	0.54470	0.35199	0.24982	0.18785	0.16246
HT-Beta	0.52113	0.22269	**0.12686**	**0.09130**	**0.07127**	**0.05927**
HT-MSE	1.45153	0.68479	0.42046	0.29258	0.21884	0.18106
HT-Rob	0.59640	0.31439	0.21479	0.15808	0.12567	0.11117
IHT	**0.26958**	**0.17517**	0.13835	0.12541	0.09784	0.08424
HA	0.09403	0.05567	0.03762	0.02964	0.02413	0.01975
HA-Beta	0.07052	0.03530	**0.02280**	**0.01655**	**0.01324**	**0.01089**
HA-MSE	**0.05811**	**0.03332**	0.02320	0.01812	0.01500	0.01266
HA-Rob	0.07085	0.04045	0.02777	0.02136	0.01752	0.01457
IHA	0.06419	0.03570	0.02536	0.02046	0.01656	0.01377
LMA	0.04282	0.02060	0.01410	0.01038	0.00850	0.00757
LMA-Trim	0.07308	0.03415	0.02154	0.01559	0.01166	0.00986
LMA-Rob	0.04226	0.01975	0.01343	0.00964	0.00784	0.00691
ILMA	**0.04176**	**0.01936**	**0.01314**	**0.00957**	**0.00767**	**0.00675**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	0.21019	0.10718	0.06967	0.05740	0.04322	0.03459
HT-Beta	0.15141	0.07191	**0.04455**	**0.03259**	**0.02473**	**0.01977**
HT-MSE	0.40684	0.17673	0.10429	0.08247	0.05855	0.04716
HT-Rob	0.13351	0.07596	0.05127	0.04275	0.03292	0.02727
IHT	**0.08246**	**0.06513**	0.04592	0.03400	0.03117	0.02523
HA	0.10994	0.06556	0.04394	0.03309	0.02765	0.02399
HA-Beta	0.08260	0.04304	**0.02782**	**0.01988**	**0.01717**	**0.01330**
HA-MSE	0.07036	**0.04162**	0.02859	0.02172	0.01910	0.01533
HA-Rob	0.07813	0.04667	0.03133	0.02386	0.02075	0.01732
IHA	**0.06669**	0.04216	0.02839	0.02068	0.01994	0.01572
LMA	0.02200	0.01127	0.00764	0.00564	0.00451	0.00421
LMA-Trim	0.04619	0.02298	0.01420	0.01092	0.00873	0.00699
LMA-Rob	0.02153	0.01069	0.00703	0.00515	0.00406	0.00364
ILMA	**0.02133**	**0.01046**	**0.00685**	**0.00497**	**0.00395**	**0.00349**

[1] For each estimation type, the minimum variance of all estimators is displayed in bold.

Table S6: Empirical variance of each estimator in the case of 100 classes.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	2.88155	1.34882	0.88706	0.69305	0.52169	0.47603
HT-Beta	1.05845	0.48952	**0.28423**	**0.19971**	**0.16152**	**0.13213**
HT-MSE	2.89600	1.32177	0.86631	0.68183	0.51872	0.46614
HT-Rob	1.37984	0.72553	0.49344	0.38777	0.30862	0.28568
IHT	**0.56117**	**0.41628**	0.29264	0.22859	0.20938	0.16932
HA	0.07893	0.04498	0.03310	0.02579	0.01965	0.01886
HA-Beta	0.06661	0.03386	0.02170	**0.01615**	**0.01198**	**0.01050**
HA-MSE	**0.05880**	**0.03195**	**0.02144**	0.01667	0.01268	0.01161
HA-Rob	0.06850	0.03754	0.02583	0.02002	0.01503	0.01401
IHA	0.06527	0.03569	0.02402	0.01862	0.01420	0.01259
LMA	0.02408	0.01088	0.00706	0.00570	0.00422	0.00349
LMA-Trim	0.08617	0.03836	0.02236	0.01624	0.01275	0.01041
LMA-Rob	0.02397	0.01070	0.00686	0.00546	0.00405	0.00328
ILMA	**0.02354**	**0.01065**	**0.00677**	**0.00538**	**0.00398**	**0.00321**
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	1.08944	0.65508	0.46603	0.45846	0.35685	0.23429
HT-Beta	0.49542	0.20477	**0.13933**	**0.09961**	**0.07672**	**0.05633**
HT-MSE	1.37562	0.83336	0.51461	0.48456	0.37648	0.26172
HT-Rob	0.56080	0.32914	0.23895	0.23332	0.17410	0.12533
IHT	**0.27049**	**0.16428**	0.14316	0.11209	0.09993	0.07674
HA	0.08611	0.04607	0.03485	0.02625	0.02284	0.01859
HA-Beta	0.07128	0.03263	**0.02196**	**0.01638**	**0.01289**	**0.01047**
HA-MSE	**0.05816**	**0.02975**	0.02229	0.01698	0.01393	0.01204
HA-Rob	0.06685	0.03453	0.02572	0.01979	0.01645	0.01382
IHA	0.06497	0.03281	0.02438	0.01842	0.01522	0.01279
LMA	0.04311	0.01950	0.01542	0.01079	0.00913	0.00735
LMA-Trim	0.07370	0.02899	0.01988	0.01459	0.01155	0.00910
LMA-Rob	**0.04224**	**0.01868**	**0.01431**	0.00991	0.00839	0.00662
ILMA	0.04242	0.01876	0.01447	**0.00976**	**0.00827**	**0.00650**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	0.49029	0.22099	0.13316	0.10439	0.08175	0.06724
HT-Beta	0.16965	0.07196	**0.04702**	**0.03481**	**0.02505**	**0.01993**
HT-MSE	1.19912	0.40963	0.23023	0.17632	0.13407	0.10501
HT-Rob	0.22692	0.11451	0.07351	0.05719	0.04625	0.03938
IHT	**0.09331**	**0.06056**	0.04969	0.03917	0.03167	0.02947
HA	0.11451	0.06434	0.04926	0.03562	0.03012	0.02468
HA-Beta	0.08208	0.04002	**0.02913**	**0.01962**	**0.01615**	**0.01213**
HA-MSE	0.06519	**0.03801**	0.03017	0.02148	0.01870	0.01549
HA-Rob	0.07881	0.04463	0.03443	0.02461	0.02130	0.01751
IHA	**0.06476**	0.03854	0.03174	0.02213	0.01932	0.01615
LMA	0.02520	0.01375	0.00926	0.00705	0.00553	0.00469
LMA-Trim	0.04389	0.02154	0.01547	0.01135	0.00973	0.00802
LMA-Rob	0.02483	0.01297	0.00849	0.00643	**0.00501**	**0.00422**
ILMA	**0.02422**	**0.01274**	**0.00837**	**0.00641**	0.00507	0.00430

[1] For each estimation type, the minimum variance of all estimators is displayed in bold.

Table S7: Empirical squared bias of each estimator under different models.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	3.25E-07	**3.26E-08**	**7.47E-06**	**1.44E-05**	**2.05E-05**	**5.05E-06**
HT-Beta	4.96E-05	6.69E-05	1.10E-02	1.13E-02	2.57E-03	2.88E-03
HT-mse	1.67E-03	2.05E-03	5.56E-04	2.55E-04	2.80E-03	2.16E-03
HT-Rob	**1.47E-09**	2.69E-07	4.53E-03	2.78E-03	8.85E-04	9.08E-04
IHT	5.72E-06	3.32E-06	5.06E-03	3.03E-03	1.42E-03	1.08E-03
LMA	8.63E-09	**1.38E-07**	**4.30E-04**	**2.16E-04**	**6.47E-04**	**3.87E-04**
LMA-Trim	1.99E-03	2.79E-03	1.02E-03	7.94E-04	2.55E-03	2.53E-03
LMA-Rob	6.77E-09	2.40E-07	7.83E-04	4.73E-04	1.07E-03	7.60E-04
ILMA	**9.40E-10**	2.04E-07	6.79E-04	3.79E-04	1.05E-03	7.08E-04
NMA_1	3.08E-05	**5.39E-09**	2.61E-03	8.38E-04	1.71E-03	4.92E-04
$INMA_1$	3.07E-05	7.07E-09	2.61E-03	8.38E-04	1.70E-03	4.94E-04
NMA_2	**7.32E-07**	1.83E-06	**3.41E-05**	**8.43E-07**	**2.01E-04**	**9.01E-05**
$INMA_2$	7.37E-07	1.82E-06	3.44E-05	9.12E-07	2.05E-04	9.49E-05
	Cycle1		Cycle4		CDF	
HT	**3.66E-04**	**1.33E-06**	**2.39E-05**	**1.57E-05**	1.38E-13	2.63E-08
HT-Beta	3.08E-02	2.85E-02	3.14E-02	3.46E-02	7.65E-09	6.25E-08
HT-mse	3.65E-03	5.74E-04	1.70E-03	1.01E-03	**1.38E-13**	**2.63E-08**
HT-Rob	1.54E-02	7.27E-03	1.27E-02	8.61E-03	3.20E-07	4.14E-08
IHT	1.24E-02	6.15E-03	1.29E-02	8.02E-03	1.04E-09	3.32E-08
LMA	1.32E-03	1.09E-03	**9.76E-05**	**7.45E-05**	**1.00E-06**	**1.44E-06**
LMA-Trim	**1.04E-04**	**4.11E-05**	1.40E-03	1.75E-03	8.45E-05	1.02E-04
LMA-Rob	2.09E-03	2.00E-03	1.37E-04	1.71E-04	1.47E-06	2.44E-06
ILMA	2.35E-03	1.97E-03	3.56E-04	2.19E-04	2.45E-06	3.02E-06
NMA_1	4.51E-03	1.27E-03	9.24E-03	1.56E-03	9.02E-10	**1.07E-11**
$INMA_1$	4.52E-03	1.28E-03	9.29E-03	1.58E-03	**8.73E-10**	1.23E-11
NMA_2	**5.99E-04**	**2.45E-04**	**7.19E-03**	1.19E-03	6.69E-07	2.25E-07
$INMA_2$	6.08E-04	2.54E-04	7.31E-03	**1.19E-03**	6.81E-07	2.33E-07

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.

[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S8: Empirical squared bias of each estimator in the case of 50 classes.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	**1.88E-08**	**1.14E-09**	**2.02E-06**	**2.57E-06**	**1.66E-06**	**4.44E-07**
HT-Beta	5.57E-05	6.27E-05	3.70E-03	3.93E-03	2.57E-04	3.61E-04
HT-mse	1.04E-03	9.86E-04	2.10E-04	1.35E-04	6.11E-04	3.75E-04
HT-Rob	4.87E-07	1.22E-07	2.22E-03	1.23E-03	2.38E-04	1.78E-04
IHT	6.67E-06	4.92E-06	1.18E-03	3.26E-04	1.03E-04	6.09E-05
LMA	3.43E-07	**5.33E-07**	**8.90E-05**	**5.33E-05**	**3.45E-05**	**1.68E-05**
LMA-Trim	2.72E-03	3.50E-03	3.53E-04	2.98E-04	4.06E-04	3.92E-04
LMA-Rob	**3.07E-07**	6.03E-07	2.61E-04	1.71E-04	1.29E-04	8.30E-05
ILMA	6.13E-07	1.12E-06	1.96E-04	9.75E-05	1.01E-04	5.20E-05
NMA_1	6.24E-06	2.80E-06	5.20E-04	9.47E-05	8.70E-05	1.58E-05
$INMA_1$	6.25E-06	2.83E-06	5.19E-04	9.47E-05	8.65E-05	1.58E-05
NMA_2	**2.11E-07**	**1.65E-06**	**1.23E-06**	**2.18E-06**	**1.84E-05**	**5.81E-06**
$INMA_2$	2.13E-07	1.68E-06	1.38E-06	2.42E-06	2.08E-05	7.44E-06
	Cycle1		Cycle4		CDF	
HT	**2.71E-05**	**1.58E-06**	2.83E-05	**6.81E-06**	4.96E-08	1.39E-08
HT-Beta	4.09E-03	3.93E-03	9.79E-04	1.46E-03	**3.21E-08**	**5.20E-09**
HT-mse	1.88E-03	4.51E-04	**2.57E-06**	1.13E-05	4.96E-08	1.39E-08
HT-Rob	4.77E-03	1.77E-03	1.50E-03	9.48E-04	7.19E-07	2.67E-07
IHT	7.53E-04	4.53E-04	2.56E-04	2.62E-05	4.30E-08	1.23E-08
LMA	**1.30E-06**	**2.98E-06**	3.50E-05	**3.55E-07**	1.01E-06	2.99E-07
LMA-Trim	2.31E-04	2.20E-04	**1.13E-06**	3.39E-05	1.46E-05	1.27E-05
LMA-Rob	1.80E-05	3.31E-05	7.27E-05	1.86E-06	1.47E-06	4.12E-07
ILMA	1.62E-05	2.69E-05	6.43E-06	1.32E-06	**9.98E-07**	**2.87E-07**
NMA_1	2.42E-04	1.20E-05	1.06E-03	**1.72E-04**	**2.16E-10**	**5.83E-11**
$INMA_1$	2.42E-04	1.20E-05	**1.06E-03**	1.73E-04	2.17E-10	6.18E-11
NMA_2	**3.69E-05**	**7.05E-06**	1.64E-03	7.61E-04	1.62E-07	1.56E-08
$INMA_2$	3.97E-05	9.20E-06	1.82E-03	8.33E-04	1.71E-07	1.78E-08

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.

[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S9: Empirical squared bias of each estimator in the case of 100 classes.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	1.28E-06	2.07E-07	**4.91E-05**	**3.28E-05**	**1.31E-06**	**4.08E-07**
HT-Beta	1.00E-04	9.81E-05	7.71E-03	7.96E-03	1.16E-03	1.22E-03
HT-mse	1.68E-03	1.36E-03	4.75E-04	6.40E-05	1.42E-03	8.33E-04
HT-Rob	**7.77E-07**	**7.71E-08**	3.55E-03	1.44E-03	5.85E-04	3.74E-04
IHT	1.34E-05	3.36E-06	2.66E-03	9.33E-04	3.82E-04	1.62E-04
LMA	6.10E-07	5.53E-08	**2.29E-04**	**5.81E-05**	**1.38E-04**	**6.03E-05**
LMA-Trim	3.51E-03	4.52E-03	6.95E-04	4.95E-04	1.31E-03	1.23E-03
LMA-Rob	**4.31E-07**	**4.78E-09**	4.78E-04	2.05E-04	3.34E-04	2.09E-04
ILMA	7.62E-07	7.62E-08	4.06E-04	1.54E-04	2.95E-04	1.44E-04
NMA_1	9.24E-07	1.59E-07	1.50E-03	1.89E-04	3.59E-04	8.17E-05
$INMA_1$	9.14E-07	**1.55E-07**	1.50E-03	1.89E-04	3.59E-04	8.16E-05
NMA_2	**7.56E-07**	2.45E-07	**1.96E-05**	**1.57E-06**	**7.39E-05**	**1.85E-05**
$INMA_2$	7.72E-07	2.80E-07	2.02E-05	1.82E-06	7.74E-05	2.07E-05
	Cycle1		Cycle4		CDF	
HT	**7.72E-07**	**2.11E-05**	**3.75E-05**	**3.39E-07**	3.76E-08	5.43E-09
HT-Beta	1.37E-02	1.68E-02	1.17E-02	1.46E-02	**1.35E-08**	**4.38E-11**
HT-mse	1.82E-03	9.85E-04	5.74E-04	7.23E-04	3.76E-08	5.43E-09
HT-Rob	6.88E-03	4.06E-03	5.52E-03	3.92E-03	7.95E-07	2.68E-07
IHT	2.87E-03	1.59E-03	3.79E-03	1.86E-03	2.97E-08	3.91E-09
LMA	**7.84E-05**	**8.51E-05**	**7.33E-05**	**9.01E-05**	4.72E-07	6.74E-08
LMA-Trim	4.01E-04	4.22E-04	1.05E-03	1.42E-03	7.00E-05	7.47E-05
LMA-Rob	2.44E-04	2.91E-04	9.05E-05	1.70E-04	6.61E-07	6.07E-08
ILMA	2.77E-04	2.45E-04	3.47E-04	3.09E-04	**2.63E-07**	**1.14E-08**
NMA_1	1.40E-03	1.28E-04	**3.90E-03**	**4.87E-04**	8.03E-10	**1.80E-10**
$INMA_1$	1.41E-03	1.28E-04	3.91E-03	4.89E-04	**7.87E-10**	1.81E-10
NMA_2	**1.16E-04**	**3.63E-05**	6.81E-03	2.21E-03	3.20E-07	3.51E-08
$INMA_2$	1.22E-04	4.00E-05	7.18E-03	2.40E-03	3.31E-07	3.96E-08

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.
[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S10: Empirical variance of each estimator under different models.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	0.00058	0.00037	0.10803	0.33272	0.06178	0.02202
HT-Beta	**0.00026**	**0.00013**	**0.01157**	**0.00677**	**0.00197**	**0.00096**
HT-mse	0.02718	0.04818	0.07315	0.19288	0.00379	0.00264
HT-Rob	0.00031	0.00019	0.03967	0.09466	0.01828	0.00732
IHT	0.00027	0.00015	0.01246	0.00945	0.00244	0.00181
LMA	0.00034	0.00019	0.00178	0.00112	0.00208	0.00155
LMA-Trim	0.00369	0.00336	**0.00114**	**0.00061**	**0.00061**	**0.00035**
LMA-Rob	0.00033	0.00018	0.00147	0.00088	0.00152	0.00104
ILMA	**0.00033**	**0.00018**	0.00130	0.00077	0.00117	0.00081
NMA_1	0.03143	0.00339	0.02606	0.01919	0.03673	0.01288
$INMA_1$	0.03143	0.00339	0.02606	0.01919	0.03674	0.01266
NMA_2	0.01320	0.00065	0.01416	0.00062	0.00421	0.00164
$INMA_2$	**0.01320**	**0.00065**	**0.01416**	**0.00062**	**0.00420**	**0.00163**
	Cycle1		Cycle4		CDF	
HT	0.17352	0.17534	0.25262	0.19871	1.03E-04	5.45E-05
HT-Beta	0.07394	**0.04311**	**0.04217**	**0.02235**	**1.03E-04**	**5.44E-05**
HT-mse	0.17370	0.17229	0.21231	0.16858	1.03E-04	5.45E-05
HT-Rob	0.10571	0.08376	0.10666	0.07991	1.06E-04	5.55E-05
IHT	**0.07218**	0.04736	0.04871	0.03620	1.03E-04	5.46E-05
LMA	0.00730	0.00547	0.01471	0.00874	7.95E-05	4.47E-05
LMA-Trim	0.01145	0.00837	**0.01201**	**0.00651**	3.35E-04	2.47E-04
LMA-Rob	0.00588	0.00403	0.01510	0.00871	7.84E-05	4.23E-05
ILMA	**0.00488**	**0.00332**	0.01387	0.00840	**7.45E-05**	**4.03E-05**
NMA_1	0.01677	0.02245	0.03109	0.03369	3.03E-07	1.56E-07
$INMA_1$	0.01665	0.02230	**0.03100**	0.03366	**3.03E-07**	**1.56E-07**
NMA_2	0.00602	0.00137	0.05723	0.02800	9.02E-06	4.62E-06
$INMA_2$	**0.00599**	**0.00135**	0.05690	**0.02792**	8.93E-06	4.58E-06

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.

[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S11: Empirical variance of each estimator in the case of 50 classes.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	0.00034	0.00020	0.02011	0.01233	0.00294	0.00173
HT-Beta	**0.00023**	**0.00013**	0.01339	**0.00790**	0.00195	**0.00108**
HT-mse	0.00975	0.00697	0.01761	0.01125	0.00178	0.00118
HT-Rob	0.00024	0.00014	0.01491	0.00990	0.00202	0.00132
IHT	0.00025	0.00014	**0.01182**	0.00908	**0.00170**	0.00120
LMA	0.00028	0.00017	0.00131	0.00073	0.00092	0.00061
LMA-Trim	0.00367	0.00311	**0.00100**	**0.00055**	**0.00056**	**0.00034**
LMA-Rob	0.00027	0.00016	0.00115	0.00064	0.00075	0.00051
ILMA	**0.00027**	**0.00016**	0.00105	0.00063	0.00070	0.00048
NMA_1	0.00052	0.00041	0.00198	0.00066	0.00065	0.00032
$INMA_1$	**0.00052**	0.00041	0.00198	0.00066	0.00065	0.00032
NMA_2	0.00094	0.00036	0.00049	0.00021	0.00058	0.00024
$INMA_2$	0.00094	**0.00036**	**0.00049**	**0.00021**	**0.00058**	**0.00023**
	Cycle1		Cycle4		CDF	
HT	0.07324	0.04545	0.03881	0.02230	1.25E-04	7.15E-05
HT-Beta	0.07685	0.04754	0.03591	0.02062	**1.25E-04**	**7.12E-05**
HT-mse	0.07466	0.04632	0.03496	0.02038	1.25E-04	7.15E-05
HT-Rob	0.07021	0.04425	0.03314	**0.01992**	1.29E-04	7.27E-05
IHT	**0.06697**	**0.04131**	**0.03172**	0.02065	1.25E-04	7.15E-05
LMA	0.00360	0.00213	0.01202	0.00649	**8.65E-05**	**4.88E-05**
LMA-Trim	0.00505	0.00302	**0.01142**	**0.00613**	1.21E-04	6.81E-05
LMA-Rob	0.00352	0.00209	0.01276	0.00677	8.94E-05	4.97E-05
ILMA	**0.00341**	**0.00199**	0.01185	0.00640	8.67E-05	4.88E-05
NMA_1	0.00343	0.00058	0.00708	0.00329	3.72E-07	1.99E-07
$INMA_1$	0.00343	0.00058	**0.00707**	**0.00329**	**3.72E-07**	**1.99E-07**
NMA_2	0.00108	0.00049	0.01847	0.00943	1.01E-05	5.58E-06
$INMA_2$	**0.00107**	**0.00048**	0.01811	0.00928	1.01E-05	5.57E-06

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.
[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S12: Empirical variance of each estimator in the case of 100 classes.

f	**0.06**	**0.10**	**0.06**	**0.10**	**0.06**	**0.10**
	Linear		Quadratic		Exponential	
HT	0.00048	0.00024	0.02890	0.01944	0.00552	0.00344
HT-Beta	**0.00025**	**0.00013**	0.01331	**0.00795**	**0.00204**	**0.00112**
HT-mse	0.02972	0.01428	0.02445	0.01694	0.00244	0.00188
HT-Rob	0.00028	0.00015	0.01875	0.01398	0.00315	0.00224
IHT	0.00026	0.00015	**0.01287**	0.00999	0.00233	0.00183
LMA	0.00031	0.00017	0.00150	0.00096	0.00141	0.00093
LMA-Trim	0.00417	0.00310	**0.00102**	**0.00059**	**0.00053**	**0.00030**
LMA-Rob	0.00030	**0.00016**	0.00127	0.00081	0.00111	0.00073
ILMA	**0.00029**	0.00016	0.00115	0.00073	0.00098	0.00068
NMA_1	0.00072	0.00063	0.00381	0.00147	0.00109	0.00065
$INMA_1$	**0.00072**	0.00063	0.00381	0.00147	0.00109	0.00065
NMA_2	0.00117	0.00043	0.00100	0.00035	0.00084	0.00036
$INMA_2$	0.00117	**0.00043**	**0.00100**	**0.00035**	**0.00084**	**0.00036**
	Cycle1		Cycle4		CDF	
HT	0.09935	0.05697	0.07821	0.04604	1.10E-04	6.06E-05
HT-Beta	0.08394	0.04646	0.04675	**0.02560**	**1.10E-04**	**6.04E-05**
HT-mse	0.10002	0.05762	0.06170	0.03819	1.10E-04	6.06E-05
HT-Rob	0.08691	0.05167	0.05624	0.03605	1.13E-04	6.16E-05
IHT	**0.07550**	**0.04619**	**0.04342**	0.03058	1.10E-04	6.06E-05
LMA	0.00486	0.00325	0.01279	0.00823	7.80E-05	4.26E-05
LMA-Trim	0.00779	0.00526	**0.01031**	**0.00632**	1.82E-04	1.06E-04
LMA-Rob	0.00448	0.00295	0.01320	0.00838	7.99E-05	4.30E-05
ILMA	**0.00412**	**0.00274**	0.01188	0.00768	**7.76E-05**	**4.21E-05**
NMA_1	0.00784	0.00199	0.01499	0.00743	3.40E-07	1.69E-07
$INMA_1$	0.00784	0.00199	**0.01498**	**0.00743**	**3.40E-07**	**1.69E-07**
NMA_2	0.00178	0.00075	0.03193	0.01649	9.86E-06	4.99E-06
$INMA_2$	**0.00176**	**0.00074**	0.03143	0.01623	9.79E-06	4.96E-06

[1] NMA_1 and $INMA_1$ are nonparametric model-assisted estimators with $h_N = 0.1$.
[2] NMA_2 and $INMA_2$ are nonparametric model-assisted estimators with $h_N = 0.25$.

Table S13: Empirical squared bias of each estimator for the "BigLucy" data set.

n_0	**256**	**512**	**853**	**1279**	**1706**	**2132**
HT	**0.0024**	**0.0000**	**0.0029**	**0.0024**	**0.0004**	**0.0000**
HT-Beta	1.4309	1.2099	1.1763	1.2734	1.2420	1.2842
HT-mse	0.1865	0.0644	0.0449	0.0324	0.0164	0.0060
HT-Rob	0.7606	0.2359	0.0942	0.0614	0.0458	0.0421
IHT	0.1800	0.0490	0.0254	0.0298	0.0220	0.0231
HA	**0.0086**	**0.0023**	**0.0051**	**0.0000**	**0.0001**	**0.0051**
HA-Beta	0.2306	0.1779	0.1993	0.1448	0.1420	0.2017
HA-mse	0.2876	0.0809	0.0494	0.0136	0.0090	0.0224
HA-Rob	0.1571	0.0516	0.0375	0.0086	0.0055	0.0181
IHA	0.0344	0.0131	0.0135	0.0024	0.0019	0.0115
LMA	**0.0095**	**0.0013**	**0.0002**	**0.0002**	**0.0011**	**0.0000**
LMA-Trim	0.0776	0.1090	0.1281	0.1364	0.1202	0.1416
LMA-Rob	0.0669	0.0229	0.0112	0.0083	0.0095	0.0036
ILMA	0.0344	0.0102	0.0058	0.0052	0.0064	0.0021

Table S14: Empirical variance of each estimator for the "BigLucy" data set.

n_0	**256**	**512**	**853**	**1279**	**1706**	**2132**
HT	28.6174	12.1969	8.6757	6.3613	4.2776	3.6931
HT-Beta	**18.7541**	**10.3609**	**6.1414**	**3.8693**	**2.9870**	**2.5393**
HT-mse	37.1760	14.3054	9.5271	7.3573	5.0335	4.1854
HT-Rob	21.9326	11.1977	7.0054	4.7089	3.4307	3.0108
IHT	19.6623	10.8806	6.5036	4.1081	3.1668	2.7076
HA	5.9140	2.8995	1.8203	1.3513	0.9600	0.8060
HA-Beta	5.0285	**2.4466**	**1.5554**	**1.0131**	**0.7110**	**0.6133**
HA-mse	**4.9454**	2.4520	1.5791	1.0383	0.7228	0.6270
HA-Rob	5.3528	2.6230	1.6651	1.1285	0.7851	0.7122
IHA	5.1299	2.5104	1.6005	1.0568	0.7298	0.6311
LMA	3.9330	2.0686	1.2485	0.9045	0.6406	0.5581
LMA-Trim	4.0394	2.1581	1.3277	1.0582	0.7738	0.6213
LMA-Rob	3.8463	2.0147	1.1639	0.8095	0.5602	0.4942
ILMA	**3.7257**	**1.9534**	**1.1190**	**0.7750**	**0.5199**	**0.4701**

S2.2 Small sample

Under the settings in Section 7.1, we consider the case of small sample in this subsection. We set the population size $N = 50000$ and the sample size $n = 5$, 10. Tables S15 - S17 show the empirical MSE of each estimator under various scenarios. It is observed that our proposed estimators (IHT, IHA and ILMA) still have the best performance in their respective estimation types, but the LMA-type estimators are no longer the best compared with the HT-type and HA-type estimators. Based on the estimated conditional bias, the robust trimmed estimators (HT-Rob, HA-Rob and LMA-Rob) perform poorly when the sample size is small. There are little differences in MSEs between the trimmed estimators (HT-Beta, HT-mse or HA-Beta, HA-mse) and the untrimmed estimators (HT or HA), this may be because the thresholds obtained by these methods only slightly modify the first-order inclusion probabilities in the case of small sample.

Table S15: Empirical MSE of each estimator under different n and $\boldsymbol{\rho}$.

n	**5**	**10**	**5**	**10**	**5**	**10**
	$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)		$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)		$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)	
HT	6.44359	6.23985	4.54939	2.46372	1.86718	0.68820
HT-Beta	6.44359	3.09681	4.54939	1.47430	1.86718	0.46111
HT-mse	6.44359	3.09681	4.54939	1.47430	1.86718	0.46111
HT-Rob	6.44359	3.09681	4.54939	1.47430	1.86718	0.46111
IHT	**1.69208**	**1.21444**	**0.77638**	**0.51936**	**0.27881**	**0.18002**
HA	0.24998	0.16019	0.29844	0.15980	0.31342	0.18828
HA-Beta	0.24998	0.15424	0.29844	0.15663	0.31342	0.18329
HA-mse	0.24998	0.15583	0.29844	0.15734	0.31342	0.18440
HA-Rob	6.44359	3.09681	4.54939	1.47430	1.86718	0.46111
IHA	**0.22230**	**0.12965**	**0.25843**	**0.13370**	**0.26550**	**0.15443**
LMA	0.31976	0.05665	0.54743	0.10453	0.26477	0.05568
LMA-Trim	1.30053	0.30703	15.31260	0.22173	17.75040	0.12268
LMA-Rob	6.44359	3.09681	4.54939	1.47430	1.86718	0.46111
ILMA	**0.31709**	**0.05558**	**0.54489**	**0.10231**	**0.26207**	**0.05504**

Table S16: Empirical MSE of each estimator in the case of 50 classes.

n	**5**	**10**	**5**	**10**	**5**	**10**
	$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)		$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)		$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)	
HT	9.23905	3.85530	3.96628	1.72830	1.67071	0.56571
HT-Beta	9.23905	2.66611	3.96628	1.19098	1.67071	0.46218
HT-mse	9.23905	2.66611	3.96628	1.19098	1.67071	0.46218
HT-Rob	9.23905	2.66611	3.96628	1.19098	1.67071	0.46218
IHT	**1.64820**	**1.15971**	**0.74807**	**0.49463**	**0.26043**	**0.16595**
HA	0.26658	0.14668	0.29336	0.16780	0.31409	0.18984
HA-Beta	0.26658	0.14397	0.29336	0.16297	0.31409	0.18524
HA-mse	0.26658	0.14458	0.29336	0.16408	0.31409	0.18621
HA-Rob	9.23905	2.66611	3.96628	1.19098	1.67071	0.46218
IHA	**0.23287**	**0.12318**	**0.24826**	**0.13606**	**0.26190**	**0.15414**
LMA	0.30426	0.05942	0.46547	0.10833	0.27972	0.06610
LMA-Trim	0.83240	0.27710	1.01280	0.22261	0.48618	0.13049
LMA-Rob	9.23905	2.66611	3.96628	1.19098	1.67071	0.46218
ILMA	**0.30315**	**0.05857**	**0.46398**	**0.10712**	**0.27930**	**0.06524**

Table S17: Empirical MSE of each estimator in the case of 100 classes.

n	**5**	**10**	**5**	**10**	**5**	**10**
	$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)		$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)		$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)	
HT	17.20283	5.77343	4.16234	2.02376	1.98476	0.89306
HT-Beta	17.20283	3.00865	4.16234	1.40449	1.98476	0.48117
HT-mse	17.20283	3.00865	4.16234	1.40449	1.98476	0.48117
HT-Rob	17.20283	3.00865	4.16234	1.40449	1.98476	0.48117
IHT	**1.69717**	**1.20437**	**0.73793**	**0.50453**	**0.26975**	**0.16662**
HA	0.25895	0.14505	0.28368	0.16739	0.31793	0.19121
HA-Beta	0.25895	0.14174	0.28368	0.16199	0.31793	0.18524
HA-mse	0.25895	0.14264	0.28368	0.16328	0.31793	0.18651
HA-Rob	17.20283	3.00865	4.16234	1.40449	1.98476	0.48117
IHA	**0.22810**	**0.12364**	**0.24019**	**0.13615**	**0.27032**	**0.15780**
LMA	0.26551	0.05757	0.46845	0.10794	0.30145	0.05662
LMA-Trim	0.80403	0.27576	0.75012	0.23485	4.45510	0.15613
LMA-Rob	17.20283	3.00865	4.16234	1.40449	1.98476	0.48117
ILMA	**0.26413**	**0.05711**	**0.46639**	**0.10606**	**0.30048**	**0.05587**

S2.3 Misspecified model

In this subsection, we consider the misspecified model based on Section 7.1. We remove the variable x_{3k} when estimating the population mean $\bar{t}_y$. In this case, the misspecified degree of model can be controlled by ρ_3, and ρ_1 describes the correlation between the research variable y_k and the first-order inclusion probability $\pi_k \propto x_{1k}$.

Tables S18 - S20 show the empirical MSE of each estimator under various scenarios. Compared with Tables 1 - 3, the efficiency of LMA-type estimators decreases. Specially, when the misspecified degree of model is high ($\rho_3 = 0.8$), the performances of LMA-type estimators are worse than those of HA-type estimators. It is found that our proposed estimators (I-HT, IHA and ILMA) perform the best overall in their respective estimation types.

Table S18: Empirical MSE of each estimator under the misspecified model.

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	3.0878	1.7346	0.9072	0.7151	0.6108	0.5085
HT-Beta	1.1982	0.6521	0.4684	0.4042	0.3686	0.3330
HT-mse	3.4093	1.9795	1.0059	0.7564	0.6746	0.5478
HT-Rob	1.5698	0.9228	0.5595	0.4451	0.3749	0.3218
IHT	**0.7638**	**0.5011**	**0.3487**	**0.2841**	**0.2442**	**0.2046**
HA	0.0933	0.0506	0.0349	0.0275	0.0240	0.0202
HA-Beta	0.0759	0.0354	0.0242	**0.0176**	**0.0141**	**0.0116**
HA-mse	**0.0643**	**0.0334**	**0.0239**	0.0182	0.0147	0.0125
HA-Rob	0.0764	0.0402	0.0280	0.0212	0.0177	0.0150
IHA	0.0729	0.0379	0.0270	0.0201	0.0163	0.0138
LMA	0.1305	0.0652	0.0455	0.0356	0.0298	0.0244
LMA-Trim	**0.0964**	**0.0415**	**0.0276**	**0.0192**	**0.0155**	**0.0126**
LMA-Rob	0.1283	0.0624	0.0429	0.0328	0.0277	0.0223
ILMA	0.1240	0.0608	0.0422	0.0312	0.0259	0.0211
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	1.6926	0.8079	0.4102	0.3424	0.3823	0.2343
HT-Beta	0.4718	0.2810	0.2168	0.1814	0.1618	0.1460
HT-mse	2.2282	1.0757	0.5399	0.4582	0.4827	0.3043
HT-Rob	0.7787	0.4196	0.2582	0.2060	0.1970	0.1404
IHT	**0.3270**	**0.2193**	**0.1647**	**0.1293**	**0.1112**	**0.0901**
HA	0.0919	0.0546	0.0376	0.0307	0.0232	0.0227
HA-Beta	0.0777	0.0413	0.0294	0.0231	0.0187	0.0167
HA-mse	0.0760	0.0429	0.0305	0.0237	0.0188	0.0165
HA-Rob	0.0783	0.0444	0.0314	0.0248	0.0191	0.0173
IHA	**0.0713**	**0.0396**	**0.0284**	**0.0219**	**0.0175**	**0.0150**
LMA	0.0655	0.0332	0.0227	0.0173	0.0149	0.0123
LMA-Trim	0.0754	0.0371	0.0264	0.0198	0.0168	0.0151
LMA-Rob	0.0633	0.0306	0.0207	0.0156	0.0131	0.0106
ILMA	**0.0608**	**0.0296**	**0.0201**	**0.0147**	**0.0127**	**0.0102**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	1.0422	0.1717	0.1759	0.0960	0.0785	0.0784
HT-Beta	0.1791	0.0901	0.0676	0.0638	0.0546	0.0499
HT-mse	3.4243	0.2897	0.5485	0.1580	0.1297	0.1814
HT-Rob	0.3888	0.1082	0.0904	0.0649	0.0522	0.0458
IHT	**0.1146**	**0.0671**	**0.0504**	**0.0442**	**0.0359**	**0.0297**
HA	0.1134	0.0643	0.0464	0.0373	0.0310	0.0256
HA-Beta	0.0991	0.0577	0.0408	0.0339	0.0287	0.0259
HA-mse	0.1088	0.0635	0.0432	0.0343	0.0274	0.0226
HA-Rob	0.0993	0.0582	0.0398	0.0315	0.0260	0.0214
IHA	**0.0850**	**0.0491**	**0.0331**	**0.0263**	**0.0220**	**0.0176**
LMA	0.0267	0.0129	0.0099	0.0075	0.0056	0.0050
LMA-Trim	0.0578	0.0336	0.0265	0.0221	0.0190	0.0186
LMA-Rob	0.0256	0.0123	0.0090	0.0068	0.0051	**0.0045**
ILMA	**0.0247**	**0.0120**	**0.0087**	**0.0066**	**0.0050**	0.0045

[1] For each estimation type, the minimum MSE of all estimators is displayed in bold.

Table S19: Empirical MSE of each estimator under the misspecified model (50 classes).

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	2.7030	1.1605	0.7319	0.5701	0.4579	0.3827
HT-Beta	1.2714	0.6555	0.4633	0.4113	0.3673	0.3266
HT-mse	2.7328	1.1925	0.7521	0.5670	0.4627	0.3865
HT-Rob	1.5182	0.7682	0.5149	0.4171	0.3462	0.2889
IHT	**0.8177**	**0.5109**	**0.3498**	**0.3056**	**0.2472**	**0.2337**
HA	0.0825	0.0500	0.0333	0.0273	0.0221	0.0181
HA-Beta	0.0705	0.0365	0.0241	**0.0176**	**0.0142**	**0.0116**
HA-mse	**0.0578**	**0.0331**	**0.0232**	0.0179	0.0149	0.0123
HA-Rob	0.0688	0.0397	0.0270	0.0213	0.0174	0.0143
IHA	0.0683	0.0394	0.0266	0.0214	0.0172	0.0149
LMA	0.1422	0.0741	0.0475	0.0396	0.0310	0.0261
LMA-Trim	**0.0849**	**0.0404**	**0.0261**	**0.0192**	**0.0149**	**0.0121**
LMA-Rob	0.1419	0.0717	0.0443	0.0367	0.0286	0.0238
ILMA	0.1351	0.0682	0.0416	0.0351	0.0268	0.0236
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	0.8644	0.5021	0.3199	0.2337	0.1836	0.1511
HT-Beta	0.4848	0.2428	0.1912	0.1558	0.1392	0.1266
HT-mse	1.0436	0.6014	0.3669	0.2651	0.2061	0.1703
HT-Rob	0.5319	0.3035	0.2173	0.1621	0.1324	0.1109
IHT	**0.3286**	**0.1887**	**0.1470**	**0.1090**	**0.1037**	**0.0835**
HA	0.0815	0.0508	0.0329	0.0254	0.0205	0.0175
HA-Beta	0.0722	0.0396	0.0260	0.0198	0.0162	0.0139
HA-mse	0.0725	0.0402	0.0268	0.0202	**0.0161**	0.0139
HA-Rob	0.0709	0.0417	0.0271	0.0211	0.0168	0.0145
IHA	**0.0680**	**0.0382**	**0.0252**	**0.0189**	0.0165	**0.0136**
LMA	0.0576	0.0312	0.0215	0.0175	0.0135	0.0113
LMA-Trim	0.0775	0.0373	0.0246	0.0187	0.0153	0.0129
LMA-Rob	0.0560	0.0283	0.0192	0.0152	0.0116	0.0095
ILMA	**0.0540**	**0.0268**	**0.0184**	**0.0140**	**0.0115**	**0.0092**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	0.3598	0.1908	0.1155	0.0746	0.0699	0.0518
HT-Beta	0.1801	0.0953	0.0729	0.0583	0.0528	0.0488
HT-mse	0.6095	0.2973	0.1672	0.1108	0.0957	0.0718
HT-Rob	0.2031	0.1182	0.0791	0.0535	0.0490	0.0392
IHT	**0.1209**	**0.0707**	**0.0535**	**0.0403**	**0.0346**	**0.0328**
HA	0.1092	0.0607	0.0398	0.0297	0.0249	0.0203
HA-Beta	0.1026	0.0547	0.0383	0.0289	0.0257	0.0219
HA-mse	0.1134	0.0594	0.0407	0.0286	0.0245	0.0196
HA-Rob	0.1003	0.0540	0.0367	0.0265	0.0229	0.0184
IHA	**0.0889**	**0.0468**	**0.0312**	**0.0227**	**0.0198**	**0.0167**
LMA	0.0240	0.0130	0.0087	0.0068	0.0052	0.0047
LMA-Trim	0.0588	0.0313	0.0225	0.0190	0.0172	0.0150
LMA-Rob	0.0234	0.0120	0.0079	0.0060	**0.0046**	**0.0041**
ILMA	**0.0225**	**0.0116**	**0.0078**	**0.0060**	0.0046	0.0043

[1] For each estimation type, the minimum MSE of all estimators is displayed in bold.

Table S20: Empirical MSE of each estimator under the misspecified model (100 classes).

f	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	**0.12**
			$\boldsymbol{\rho}$=(0.1, 0.3, 0.8)			
HT	2.4691	1.6945	1.1882	0.8149	0.5627	0.4477
HT-Beta	1.1076	0.5436	0.4148	0.3508	0.3279	0.2911
HT-mse	2.4915	1.6849	1.1632	0.7900	0.5696	0.4502
HT-Rob	1.3470	0.8960	0.5823	0.4346	0.3378	0.2663
IHT	**0.7210**	**0.4150**	**0.3129**	**0.2525**	**0.2184**	**0.1818**
HA	0.0851	0.0486	0.0340	0.0253	0.0211	0.0184
HA-Beta	0.0696	0.0351	0.0224	**0.0162**	**0.0130**	**0.0115**
HA-mse	**0.0596**	**0.0330**	**0.0220**	0.0166	0.0140	0.0125
HA-Rob	0.0707	0.0382	0.0262	0.0196	0.0163	0.0143
IHA	0.0665	0.0372	0.0253	0.0191	0.0163	0.0151
LMA	0.1555	0.0770	0.0501	0.0390	0.0313	0.0262
LMA-Trim	**0.0861**	**0.0393**	**0.0248**	**0.0179**	**0.0145**	**0.0128**
LMA-Rob	0.1541	0.0748	0.0462	0.0360	0.0286	0.0236
ILMA	0.1478	0.0705	0.0429	0.0333	0.0272	0.0228
			$\boldsymbol{\rho}$=(0.5, 0.4, 0.3)			
HT	1.0412	0.4589	0.3347	0.2421	0.1913	0.1614
HT-Beta	0.5105	0.2540	0.2007	0.1616	0.1430	0.1288
HT-mse	1.4195	0.5786	0.3954	0.2795	0.2270	0.1867
HT-Rob	0.5947	0.2993	0.2296	0.1759	0.1414	0.1188
IHT	**0.3349**	**0.2055**	**0.1630**	**0.1210**	**0.1041**	**0.0881**
HA	0.1005	0.0540	0.0379	0.0283	0.0248	0.0200
HA-Beta	0.0828	0.0409	0.0299	0.0235	0.0189	0.0162
HA-mse	0.0821	0.0421	0.0311	0.0241	0.0193	0.0160
HA-Rob	0.0849	0.0441	0.0317	0.0243	0.0201	0.0166
IHA	**0.0767**	**0.0402**	**0.0295**	**0.0224**	**0.0185**	**0.0153**
LMA	0.0558	0.0295	0.0205	0.0158	0.0128	0.0107
LMA-Trim	0.0878	0.0383	0.0280	0.0216	0.0171	0.0151
LMA-Rob	0.0538	0.0274	0.0187	0.0142	0.0115	0.0094
ILMA	**0.0513**	**0.0267**	**0.0180**	**0.0136**	**0.0115**	**0.0093**
			$\boldsymbol{\rho}$=(0.8, 0.3, 0.1)			
HT	0.3225	0.1541	0.1107	0.0780	0.0679	0.0523
HT-Beta	0.1610	0.0850	0.0626	0.0545	0.0461	0.0430
HT-mse	0.6303	0.2821	0.1991	0.1220	0.1135	0.0845
HT-Rob	0.1858	0.0986	0.0702	0.0535	0.0444	0.0364
IHT	**0.1077**	**0.0656**	**0.0467**	**0.0387**	**0.0309**	**0.0272**
HA	0.1078	0.0605	0.0438	0.0318	0.0264	0.0222
HA-Beta	0.0989	0.0537	0.0394	0.0317	0.0259	0.0235
HA-mse	0.1106	0.0583	0.0413	0.0315	0.0240	0.0210
HA-Rob	0.0988	0.0539	0.0384	0.0291	0.0226	0.0194
IHA	**0.0848**	**0.0455**	**0.0321**	**0.0245**	**0.0192**	**0.0166**
LMA	0.0235	0.0121	0.0079	0.0060	0.0047	0.0040
LMA-Trim	0.0601	0.0315	0.0254	0.0200	0.0184	0.0168
LMA-Rob	0.0225	0.0114	**0.0072**	**0.0055**	**0.0042**	**0.0036**
ILMA	**0.0224**	**0.0114**	0.0073	0.0056	0.0042	0.0037

[1] For each estimation type, the minimum MSE of all estimators is displayed in bold.

S2.4 LMA vs. ILMA

In this subsection, we compare the empirical performances of the LMA and ILMA estimators in terms of their MSEs, biases, variances and coverage rates. Based on πps sampling, $D = 2000$ replicated samples are drawn from the finite population given in Section 7.1, and then the squared-bias (Bias^2), variance (Var), mean squared error (MSE), and coverage rate (CR) of each estimator are computed empirically. Denote K* as the threshold of ILMA estimator, and $\mathrm{Re} = \frac{|\mathrm{MSE1}-\mathrm{MSE2}|}{|\mathrm{MSE1}|} \times 100\%$, where MSE1 and MSE2 denote the empirical MSEs of the traditional estimator and its improved estimator, respectively. CR values are calculated based on 95 percent standard normal confidence interval: $\hat{\theta} \pm 1.96 \times \sqrt{\widehat{\mathrm{MSE}}(\hat{\theta})}$.

From Tables S21 - S23, the improved estimator has smaller MSE than the traditional estimator, and the threshold K^* decreases as the sample size increases. Under the same sample size, the MSEs in the case $\boldsymbol{\rho} = (0.5, 0.4, 0.3)$ are higher than those in other cases, $\boldsymbol{\rho} = (0.1, 0.3, 0.8)$ and $\boldsymbol{\rho} = (0.8, 0.3, 0.1)$. This may be due to their different SNRs (Signal Noise Ratios). Additionally, the coverage rates CR1 and CR2 are roughly the same under various scenarios with the latter corresponding to a shorter interval length, and both grow when n increases.

Table S21: Empirical performances of $\hat{\hat{t}}_{yr}$ and $\hat{\hat{t}}^*_{yr}$ for different n and $\boldsymbol{\rho}$.

n	MSE1	MSE2	Bias^{2}1	Bias^{2}2	Var1	Var2	CR1	CR2	K*	Re
					$\boldsymbol{\rho} = (0.1, 0.3, 0.8)$					
20	0.02379	0.02334	4.30E-05	4.20E-05	0.02375	0.02330	78%	78%	151	2%
40	0.01064	0.01050	8.55E-06	6.90E-06	0.01063	0.01049	85%	85%	106	1%
60	0.00698	0.00681	3.92E-06	2.96E-06	0.00698	0.00680	88%	87%	83	3%
80	0.00535	0.00514	4.44E-07	6.56E-07	0.00535	0.00514	90%	89%	73	4%
100	0.00458	0.00434	2.76E-06	1.46E-06	0.00458	0.00433	88%	88%	64	5%
120	0.00358	0.00344	7.53E-07	3.38E-07	0.00358	0.00344	90%	89%	58	4%
150	0.00289	0.00277	2.16E-06	6.68E-07	0.00289	0.00277	90%	90%	51	4%
					$\boldsymbol{\rho} = (0.5, 0.4, 0.3)$					
20	0.05074	0.04989	3.03E-05	4.09E-05	0.05071	0.04985	77%	77%	157	2%
40	0.02328	0.02204	1.15E-06	3.38E-07	0.02328	0.02204	85%	85%	104	5%
60	0.01424	0.01319	2.41E-05	3.03E-05	0.01421	0.01316	89%	88%	83	7%
80	0.01110	0.01031	7.58E-09	5.24E-08	0.01110	0.01031	88%	88%	71	7%
100	0.00926	0.00852	1.59E-05	1.46E-05	0.00924	0.00850	89%	89%	65	8%
120	0.00745	0.00679	3.12E-09	5.87E-07	0.00745	0.00679	90%	90%	60	9%
150	0.00592	0.00534	7.03E-06	1.42E-05	0.00591	0.00532	91%	90%	52	10%
					$\boldsymbol{\rho} = (0.8, 0.3, 0.1)$					
20	0.02371	0.02318	6.17E-05	5.72E-05	0.02365	0.02313	80%	79%	158	2%
40	0.01118	0.01086	1.34E-07	1.07E-06	0.01118	0.01085	86%	86%	108	3%
60	0.00715	0.00685	3.93E-06	3.91E-06	0.00715	0.00685	88%	88%	86	4%
80	0.00564	0.00527	6.52E-07	3.04E-07	0.00564	0.00527	88%	88%	74	7%
100	0.00421	0.00392	1.52E-06	1.21E-06	0.00421	0.00392	90%	90%	67	7%
120	0.00371	0.00330	5.47E-06	5.55E-06	0.00370	0.00329	90%	90%	61	11%
150	0.00297	0.00265	2.02E-05	1.37E-05	0.00295	0.00264	90%	90%	54	11%

Table S22: Empirical performances of $\hat{\hat{t}}_{yr}$ and $\hat{\hat{t}}^{*}_{yr}$ in the case of 50 classes.

n	MSE1	MSE2	$\text{Bias}^2 1$	$\text{Bias}^2 2$	Var1	Var2	CR1	CR2	K^*	Re
					$\boldsymbol{\rho} = (0.1, 0.3, 0.8)$					
20	0.02322	0.02252	7.45E-06	1.19E-05	0.02322	0.02251	77%	76%	160	3%
40	0.01224	0.01157	3.47E-05	3.18E-05	0.01220	0.01154	82%	82%	119	5%
60	0.00770	0.00731	1.14E-05	1.17E-05	0.00769	0.00730	85%	85%	100	5%
80	0.00679	0.00620	6.27E-06	2.02E-06	0.00678	0.00620	86%	86%	80	9%
100	0.00497	0.00456	1.62E-06	4.15E-06	0.00497	0.00455	89%	89%	70	8%
120	0.00410	0.00366	5.85E-06	1.13E-05	0.00410	0.00365	89%	90%	60	11%
150	0.00354	0.00309	1.78E-07	2.60E-06	0.00354	0.00309	90%	90%	60	13%
					$\boldsymbol{\rho} = (0.5, 0.4, 0.3)$					
20	0.05210	0.05036	1.64E-04	1.80E-04	0.05193	0.05018	76%	76%	160	3%
40	0.02577	0.02450	2.81E-05	4.03E-05	0.02574	0.02446	83%	83%	109	5%
60	0.01692	0.01580	4.62E-06	1.26E-05	0.01692	0.01579	85%	86%	89	7%
80	0.01429	0.01349	2.24E-06	5.51E-06	0.01429	0.01348	86%	86%	80	6%
100	0.01078	0.01001	1.75E-07	2.11E-06	0.01078	0.01000	87%	87%	66	7%
120	0.00906	0.00875	2.88E-06	4.47E-06	0.00906	0.00875	88%	88%	60	3%
150	0.00713	0.00681	2.35E-05	2.99E-05	0.00711	0.00678	89%	88%	59	4%
					$\boldsymbol{\rho} = (0.8, 0.3, 0.1)$					
20	0.02260	0.02198	9.71E-05	1.04E-04	0.02250	0.02187	79%	78%	147	3%
40	0.01070	0.01038	9.74E-06	1.08E-05	0.01069	0.01037	83%	83%	100	3%
60	0.00753	0.00703	6.01E-06	1.22E-05	0.00753	0.00701	86%	86%	83	7%
80	0.00546	0.00508	2.20E-06	7.81E-06	0.00546	0.00507	88%	88%	80	7%
100	0.00434	0.00389	4.43E-06	1.09E-05	0.00434	0.00388	90%	90%	64	10%
120	0.00380	0.00349	4.74E-06	5.99E-06	0.00380	0.00349	90%	90%	60	8%
150	0.00294	0.00266	1.54E-05	1.97E-05	0.00293	0.00264	90%	90%	60	10%

Table S23: Empirical performances of $\hat{\hat{t}}_{yr}$ and $\hat{\hat{t}}^*_{yr}$ in the case of 100 classes.

n	MSE1	MSE2	Bias^{2}1	Bias^{2}2	Var1	Var2	CR1	CR2	K*	Re
					$\boldsymbol{\rho} = (0.1, 0.3, 0.8)$					
20	0.02464	0.02365	1.30E-07	2.94E-06	0.02464	0.02365	78%	78%	150	4%
40	0.01206	0.01141	1.28E-05	1.04E-05	0.01204	0.01140	85%	85%	109	5%
60	0.00823	0.00761	1.23E-05	7.34E-06	0.00822	0.00761	87%	87%	88	8%
80	0.00622	0.00573	2.00E-06	5.56E-06	0.00622	0.00572	88%	88%	71	8%
100	0.00493	0.00441	9.00E-07	6.98E-06	0.00493	0.00440	90%	89%	65	11%
120	0.00436	0.00381	4.90E-07	5.41E-08	0.00436	0.00381	90%	90%	60	13%
150	0.00342	0.00298	2.39E-07	2.72E-06	0.00342	0.00298	91%	90%	52	13%
					$\boldsymbol{\rho} = (0.5, 0.4, 0.3)$					
20	0.04461	0.04313	1.11E-06	2.34E-06	0.04460	0.04312	79%	78%	143	3%
40	0.02180	0.02038	7.79E-06	7.37E-06	0.02180	0.02037	84%	84%	100	7%
60	0.01420	0.01347	2.32E-05	2.32E-05	0.01418	0.01345	86%	86%	80	5%
80	0.01057	0.00984	3.73E-06	2.97E-06	0.01057	0.00984	89%	88%	70	7%
100	0.00906	0.00805	7.67E-07	7.81E-07	0.00905	0.00805	89%	89%	61	11%
120	0.00730	0.00672	5.29E-06	5.55E-06	0.00729	0.00672	90%	90%	59	8%
150	0.00568	0.00521	3.81E-07	1.49E-07	0.00568	0.00521	91%	90%	50	8%
					$\boldsymbol{\rho} = (0.8, 0.3, 0.1)$					
20	0.02251	0.02194	5.06E-07	2.74E-08	0.02251	0.02194	80%	80%	151	3%
40	0.01183	0.01141	3.03E-06	5.76E-07	0.01183	0.01141	84%	84%	110	4%
60	0.00767	0.00734	7.38E-07	1.47E-06	0.00767	0.00734	88%	87%	89	4%
80	0.00588	0.00564	6.87E-06	1.02E-05	0.00587	0.00563	88%	88%	76	4%
100	0.00472	0.00453	1.01E-05	1.36E-05	0.00471	0.00452	89%	88%	67	4%
120	0.00383	0.00370	8.72E-07	1.46E-06	0.00383	0.00370	89%	89%	60	3%
150	0.00294	0.00278	1.81E-06	3.36E-06	0.00294	0.00278	90%	90%	56	6%

S2.5 Other supplements

In this subsection, we report some numerical characteristics and histograms of the first-order inclusion probabilities given in Section 7.3. The threshold K^* (= 1800, 1080, 756, 576, 504, 451) obtained by our method decreases as n_0 increases. Table S24 shows the mean, standard deviation, five numbers and cut-point of the first-order inclusion probabilities under different n_0. Figure 1 shows the histograms of the first-order inclusion probabilities under different n_0.

Table S24: Numerical characteristics of the first-order inclusion probabilities

n_0	mean	sd	min	lower-hinge	median	upper-hinge	max	cut-point
256	0.003001	0.001858	0.000007	0.001607	0.002711	0.003983	0.017540	0.000552
512	0.006003	0.003716	0.000014	0.003214	0.005423	0.007966	0.035079	0.000922
853	0.010000	0.006191	0.000023	0.005355	0.009034	0.013272	0.058443	0.001281
1279	0.014995	0.009283	0.000035	0.008030	0.013546	0.019900	0.087630	0.001676
1706	0.020001	0.012382	0.000047	0.010711	0.018068	0.026544	0.116885	0.001956
2132	0.024995	0.015474	0.000058	0.013385	0.022580	0.033172	0.146072	0.002211

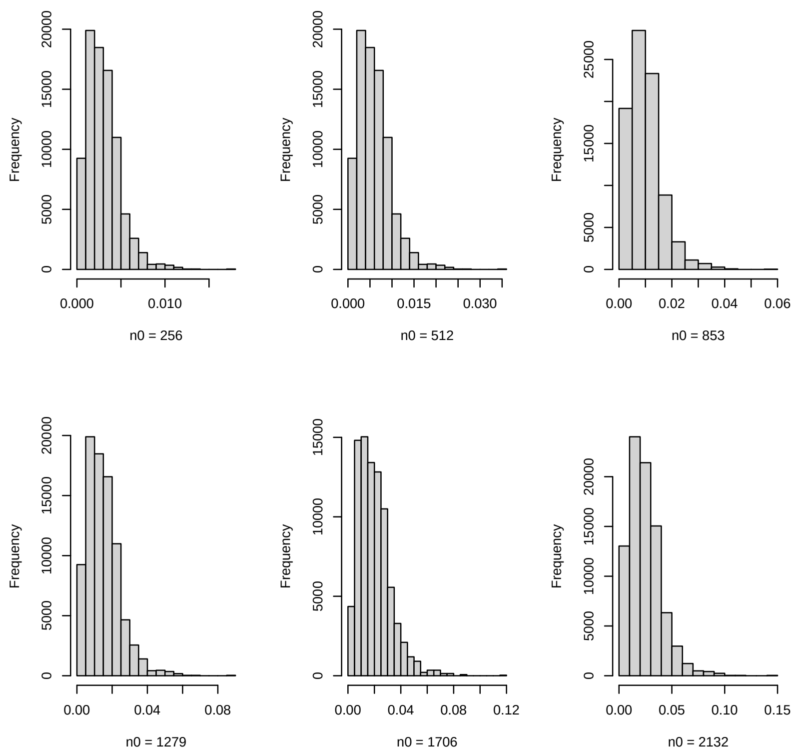

Figure 1: Histograms of the first-order inclusion probabilities

Bibliography

Breidt, F. J. and J. D. Opsomer (2000). Local polynomial regression estimators in survey sampling. *Annals of Statistics 28*, 1026–1053.

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

Graybill, F. A. (1983). *Matrices with Applications in Statistics (Second Edition)*. Belmont, CA: Wadsworth.

Särndal, C. E., B. Swensson, and J. H. Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer.

Zong, X., R. Zhu, and G. Zou (2019). Improved Horvitz-Thompson estimator in survey sampling. *Survey Methodology 45*, 165–184.