Supplemental materials for "Assessing Heterogeneous Causal Effects across Clusters in Partially Nested Designs"

Supplemental materials: Identification result

In what follows, we first identify the effect ATE_k under no-interference and then identify the effect ATE_k^{cl} under the within-cluster interference.

Under no-interference

The effect ATE_k defined under no-interference can be written as:

$ATE_k = \mathrm{E}[Y_i(1) \mid K_i(1) = k] - \mathrm{E}[Y_i(0) \mid K_i(1) = k] = \frac{\mathrm{E}[Y_i(1)\mathbf{1}(K_i(1)=k)] - \mathrm{E}[Y_i(0)\mathbf{1}(K_i(1)=k)]}{\mathrm{E}[\mathbf{1}(K_i(1)=k)]}$.

For the denominator, under the treatment ignorability assumption, $K_i(1) \perp\!\!\!\perp T_i \mid X_i$, we can write the potential cluster assignment probability as:

$$\mathrm{E}[\mathbf{1}(K_i(1) = k)] = \mathrm{E}[\mathrm{E}[K_i(1) = k \mid X_i]] = \mathrm{E}[\mathrm{E}[K_i(1) = k \mid T_i = 1, X_i]] \quad (1)$$

$$= \mathrm{E}[\mathrm{E}[K_i = k \mid T_i = 1, X_i]] = \mathrm{E}[p_k(X_i)]. \quad (2)$$

For the numerator, we first consider the potential control outcome. Under the treatment ignorability assumption $Y_i(0) \perp\!\!\!\perp T_i \mid X_i$ and the principal ignorability assumption $Y_i(0) \perp\!\!\!\perp K_i(1) \mid X_i$, we have:

$$\begin{aligned}
\mathrm{E}\left[Y_i(0)\mathbf{1}(K_i(1) = k)\right] &= \mathrm{E}\left[\mathrm{E}[Y_i(0)\mathbf{1}(K_i(1) = k)|X_i]\right] \\
&= \mathrm{E}\left[\mathrm{E}[Y_i(0)|X_i] \cdot \mathrm{E}[K_i(1)|X_i]\right] \\
&= \mathrm{E}\left[\mathrm{E}[Y_i(0)|T_i = 0, X_i] \cdot \mathrm{E}[K_i(1)|T_i = 1, X_i]\right] \\
&= \mathrm{E}\left[\mathrm{E}[Y_i|T_i = 0, X_i] \cdot \mathrm{E}[K_i|T_i = 1, X_i]\right] = \mathrm{E}[\mu_{y|0}(X_i)p_k(X_i)].
\end{aligned}$$

For the potential treatment outcome, under no-interference and the treatment ignorability assumption IA1 (i) $T_i \perp\!\!\!\perp K_i(1) \mid X_i$ and (ii) $Y_i(1) \perp\!\!\!\perp T_i \mid X_i$, we have

$$
\begin{aligned}
\mathrm{E}\left[Y_i(1)\mathbf{1}(K_i(1)=k)\right] &= \mathrm{E}\left[\mathrm{E}[Y_i(1)\mathbf{1}(K_i(1)=k) \mid X_i]\right] && (3)\\
&= \mathrm{E}\left[\mathrm{E}[Y_i(1)\mathbf{1}(K_i(1)=k) \mid T_i=1, X_i]\right] && (4)\\
&= \mathrm{E}\left[\mathrm{E}[Y_i\mathbf{1}(K_i(1)=k) \mid T_i=1, X_i]\right] && (5)\\
&= \mathrm{E}\left[\mathrm{E}[Y_i(1) \mid K_i(1)=k, T_i=1, X_i]\mathrm{p}(K_i(1)=k \mid T_i=1, X_i)\right] && (6)\\
&= \mathrm{E}\left[\mathrm{E}[Y_i \mid K_i=k, T_i=1, X_i]\mathrm{p}(K_i=k \mid T_i=1, X_i)\right] && (7)\\
&= \mathrm{E}[\mu_{y|1,k}(X_i)p_k(X_i)]. && (8)
\end{aligned}
$$

Putting the numerator and denominator together, we have the identification result in the main text.

Under the within-cluster interference

We can write the effect definition ATE_k^{cl} as $ATE_k^{\mathrm{cl}} = \mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid K_i(1)=k] - \mathrm{E}[Y_i(0) \mid K_i(1)=k]$. In the above, we have identified the cluster mean potential control outcome $\mathrm{E}[Y_i(0) \mid K_i(1)=k]$, for which there is no interference. Therefore, the task to identify ATE_k^{cl} becomes the task of identifying $\mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid K_i(1)=k]$, which we describe here.

By iterated expectations, we can write:

$$
\mathrm{E}\left[Y_i(1, \mathbf{T}_{-ik}) \mid K_i(1)=k\right] =
$$
$$
\sum_{x_i} \mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid X_i = x_i, K_i(1)=k]\mathrm{p}(X_i = x_i \mid K_i(1)=k),
$$

where $\sum_{x_i}$ is summing over (or integrating over) all possible values of x_i.

Under the treatment ignorability assumption IA1(ii'), we can write the conditional mean of the potential outcome

$$\mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid X_i = x_i, K_i(1) = k] \tag{9}$$

$$= \mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid T_i = 1, X_i = x_i, K_i(1) = k] \tag{10}$$

$$= \mathrm{E}[Y_i \mid T_i = 1, X_i = x_i, K_i(1) = k] \tag{11}$$

$$= \mathrm{E}[Y_i \mid T_i = 1, X_i = x_i, K_i = k], \tag{12}$$

which is $\mu_{y|1,k}(X_i = x_i)$.

For $\mathrm{p}(X_i = x_i \mid K_i(1) = k)$, we can re-write it under IA1 (i) (i.e., as in Eq. 1) and the Bayes rule, $\mathrm{p}(X_i = x_i \mid K_i(1) = k) = \frac{\mathrm{p}(K_i(1)=k|X_i=x_i)\mathrm{p}(X_i=x_i)}{\mathrm{p}(K_i(1)=k)} = \frac{\mathrm{p}(K_i=k|T_i=1,X_i=x_i)\mathrm{p}(X_i=x_i)}{\mathrm{E}[p_k(X_i)]}$, where in the numerator $\mathrm{p}(K_i = k \mid T_i = 1, X_i = x_i)$ is $p_k(X_i = x_i)$. Putting the above together, we have

$$\mathrm{E}\left[Y_i(1, \mathbf{T}_{-ik}) \mid K_i(1) = k\right] = \\ = \sum_{x_i} \mu_{y|1,k}(X_i = x_i) \frac{p_k(X_i = x_i)}{\mathrm{E}[p_k(X_i)]} \mathrm{p}(X_i = x_i) = \frac{\mathrm{E}[\mu_{y|1,k}(X_i) p_k(X_i)]}{\mathrm{E}[p_k(X_i)]}, \tag{13}$$

which is the identification result in the main text.

Simulation Study: Supplemental Results

The simulation results for homogeneous cluster-specific treatment effects are shown in Figures S1, S2, and S3.

Similar to the patterns described in the main text, in Figure S1, the unadjusted estimator (i.e., simple difference in means between each cluster k and the control arm) was biased even when the treatment assignment was randomized, unless the cluster assignment was also randomized. The developed estimator ($\hat{\psi}_{ATE_k}$) yielded low bias across different combinations of nonrandomized/randomized treatment and cluster assignments.

In Figure S2 for the confidence intervals, the bootstrap-based interval method performed more satisfactorily than the Wald-type interval method. The Wald-type intervals showed undercoverage issues with $J = 16$ clusters in the treatment arm and sample size $n = 400$, under nonrandomized assignments for both the treatment and cluster assignments. The bootstrap-based intervals were more conservative than the Wald-type intervals.

For the MSE (Figure S3), both the total sample size n and the number of clusters J influenced the MSE. Holding constant the total sample size, the MSEs were generally lower with fewer clusters (or larger cluster sizes) than with more clusters (or smaller cluster sizes).

Sensitivity Analysis

We conduct a sensitivity analysis for the principal ignorability assumption (IA2). Our sensitivity analysis method is based on methods in the principal score literature (Ding & Lu, 2017; Jiang et al., 2022; Nguyen et al., 2023), with adaptations to suit the PND data structure. For notation simplicity, we omit subscript i for individual i when there is no confusion.

Sensitivity Analysis Method

To describe the sensitivity analysis method, we re-write the principal ignorability assumption as: $\mathrm{E}[Y(0) \mid K(1) = k, X] = \mathrm{E}[Y(0) \mid K(1) \neq k, X] = \mathrm{E}[Y(0) \mid X]$ for all $k = 1, ..., J$.

Then, the sensitivity analysis proceeds as follows. (1) The first step is to specify a degree of deviation from this assumption, with the "degree of deviation" quantified by a sensitivity parameter. (2) The second step is to modify the identification result ψ_{ATE_k} for the cluster-specific treatment effects (ATE_k and ATE_k^{cl}) under this specified sensitivity parameter. (3) Finally, the third step is to estimate the modified identification result, given a user-specified value of the sensitivity parameter. If the estimation results remain similar to the original results (i.e., the results assuming the principal ignorability), then this would indicate that the estimation results are robust to the specified degree of violation of the principal ignorability assumption.

For the first step, we use a mean difference as the sensitivity parameter, considering that the outcome in the example is continuous and the magnitude of a mean difference is intuitive to interpret. Specifically, we specify the deviation as: $\mathrm{E}[Y(0) \mid K(1) = k, X] - \mathrm{E}[Y(0) \mid K(1) \neq k, X] = \mathsf{diff}_k$. diff_k is the sensitivity parameter, and the value is specified by researchers. For example, we may specify it based on standardized mean differences (e.g., medium size 0.3), such as specifying $\mathsf{diff}_k = 0.3 s_y$, where s_y is the sample standard deviation of the outcome of the control arm.

For the second step, we first note that the following equality holds: $p_k(X)\mathrm{E}[Y(0) \mid K(1) = k, X] + [1 - p_k(X)]\mathrm{E}[Y(0) \mid K(1) \neq k, X] = \mathrm{E}[Y(0) \mid X] = \mu_{y|0}(X)$. The second equal sign follows from the treatment assignment ignorability assumption. Then, by plugging in the sensitivity parameter, we have $p_k(X)\mathrm{E}[Y(0) \mid K(1) = k, X] + [1 - p_k(X)]\{\mathrm{E}[Y(0) \mid K(1) = k, X] - \mathsf{diff}_k\} = \mu_{y|0}(X)$, which leads to $\mathrm{E}[Y(0) \mid K(1) = k, X] = \mu_{y|0}(X) + \mathsf{diff}_k[1 - p_k(X)]$. Using this in the identification, we have the modified identification result for $\mathrm{E}[Y(0) \mid K(1) = k]$ as follows:

$$\begin{aligned}\mathrm{E}[Y(0) \mid K(1) = k] &= \frac{\mathrm{E}[\{\mu_{y|0}(X) + \mathsf{diff}_k[1 - p_k(X)]\}p_k(X)]}{\mathrm{E}[p_k(X)]} \\ &= \frac{\mathrm{E}[\mu_{y|0}(X)p_k(X)]}{\mathrm{E}[p_k(X)]} + \mathsf{diff}_k - \mathsf{diff}_k\frac{\mathrm{E}[p_k(X)p_k(X)]}{\mathrm{E}[p_k(X)]} \end{aligned} \tag{14}$$

where the first term is the identification result under the principal ignorability assumption; the second and third terms quantify the influence of the deviation from this assumption. Note that for the cluster-mean potential treatment outcomes $\mathrm{E}[Y_i(1) \mid K_i(1) = k]$ and $\mathrm{E}[Y_i(1, \mathbf{T}_{-ik}) \mid K_i(1) = \mathbf{K}_{-ik}(1) = k]$, the identification does not require the principal ignorability assumption, and therefore the identification results for them remains unchanged. Hence, the modified identification result for the cluster-specific treatment effect (ATE_k or ATE_k^{cl}) is:

$$\psi_{ATE_k}(\mathsf{diff}_k) = \psi_{ATE_k} + \mathsf{diff}_k - \mathsf{diff}_k\frac{\mathrm{E}[p_k(X)p_k(X)]}{\mathrm{E}[p_k(X)]}, \tag{15}$$

which is a function of the user-specified value of the sensitivity parameter.

For the third step, the only additional term we need to estimate in the modified identification result is the ratio in the third term, which we denote as $\frac{\mathrm{E}[p_k(X)p_k(X)]}{\mathrm{E}[p_k(X)]} := \psi_{k,\text{sens}}$. To estimate this, we estimate the denominator $\mathrm{E}[p_k(X)]$ using the same sample average of the doubly robust estimator, $\mathrm{E}[p_k(X)] = \frac{1}{n}\sum_{i=1}^{n} \hat{p}_k^{\mathsf{dr}}(O_i)$. For the numerator $\mathrm{E}[p_k(X)p_k(X)]$, we use a robust estimator from Kennedy (2023), which is the sample average $\mathrm{E}[p_k(X)p_k(X)] = \frac{1}{n}\sum_{i=1}^{n} \hat{\phi}_k^{\mathsf{sens}}(X_i)$, where $\hat{\phi}_k^{\mathsf{sens}}(X_i) = 2\hat{p}_k(X_i)\frac{1(T_i=1)}{\hat{\pi}_t(X_i)}[1(K_i = k) - \hat{p}_k(X_i)] + [\hat{p}_k(X_i)]^2$. Combining the denominator and numerator, we have the estimator for the modified identification result as:

$$\begin{aligned}\hat{\psi}_{ATE_k}(\mathsf{diff}_k) &= \hat{\psi}_{ATE_k} + \mathsf{diff}_k - \mathsf{diff}_k \frac{\sum_{i=1}^{n} \hat{\phi}_k^{\mathsf{sens}}(X_i)/n}{\sum_{i=1}^{n} \hat{p}_k^{\mathsf{dr}}(O_i)/n} \\ &= \frac{\frac{1}{n}\sum_{i=1}^{n} [\hat{\phi}_{1k,\text{num}}(O_i) - \hat{\phi}_{0k,\text{num}}(O_i) - \mathsf{diff}_k \sum_{i=1}^{n} \hat{\phi}_k^{\mathsf{sens}}(X_i)]}{\frac{1}{n}\sum_{i=1}^{n} \hat{p}_k^{\mathsf{dr}}(O_i)} + \mathsf{diff}_k \end{aligned} \tag{16}$$

We apply the developed bootstrap-based method to obtain a 95% interval for the effect estimate $\hat{\psi}_{ATE_k}(\mathsf{diff}_k)$ of each cluster under the sensitivity parameter.

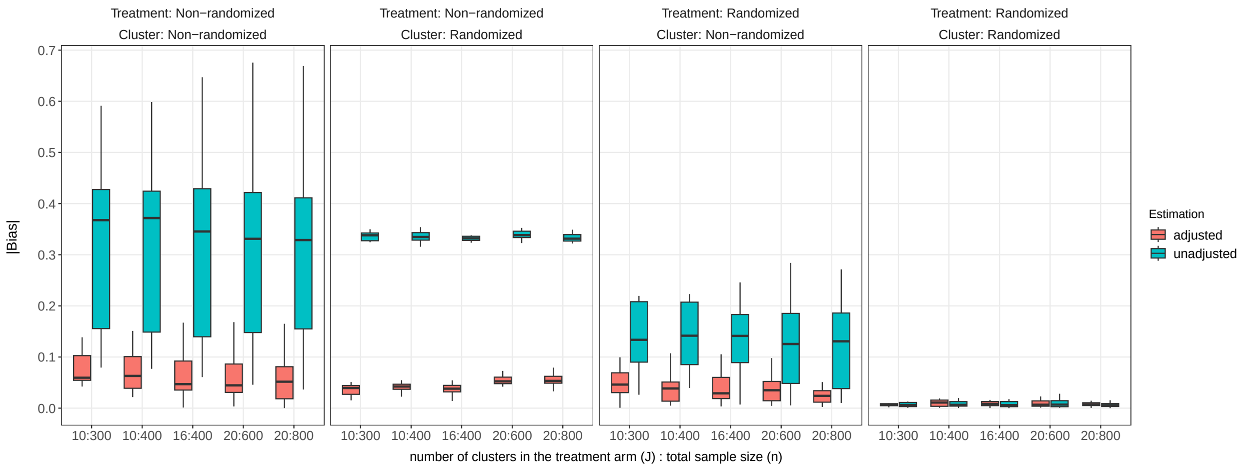
Sensitivity Analysis for the Example

Implementing this sensitivity analysis method, we specified a range of values for the mean difference sensitivity parameter diff_k based on a standardized mean difference, diff_k/s_y = $-0.3, -0.1, 0.1$ and 0.3, where s_y is the sample standard deviation of the outcome of the control arm. Figure S4 shows the sensitivity analysis results.

From the results, the estimation results for the cluster-specific treatment effects in cluster $k = 1, 2$ were relatively insensitive to the deviations from the principal ignorability assumption, because the estimates and intervals remained below zero. The estimation result for cluster $k = 3$ was sensitive to the deviation from the principal ignorability assumption quantified by the mean difference sensitivity parameter value $\mathsf{diff}_k = 0.3s_y$, because under this degree of deviation the 95% interval changed from covering zero to below zero; the result was insensitive to the other degrees of deviation. Following similar interpretations of the sensitivity results, the estimation results for cluster $k = 8, 9, 10$ were sensitive to the deviations quantified by the mean difference sensitivity parameter value $\mathsf{diff}_k = -0.3s_y, -0.1s_y$, as under this degree of deviation, the 95% intervals changed from covering zero to above zero. For cluster $k = 11, 12$, the estimation results were sensitive to the deviation quantified by the mean difference sensitivity parameter value $\mathsf{diff}_k = 0.3s_y$, as their 95% intervals changed from above zero to covering zero under this degree of deviation; the results were insensitive to the other examined degrees of deviation.

Figure S1

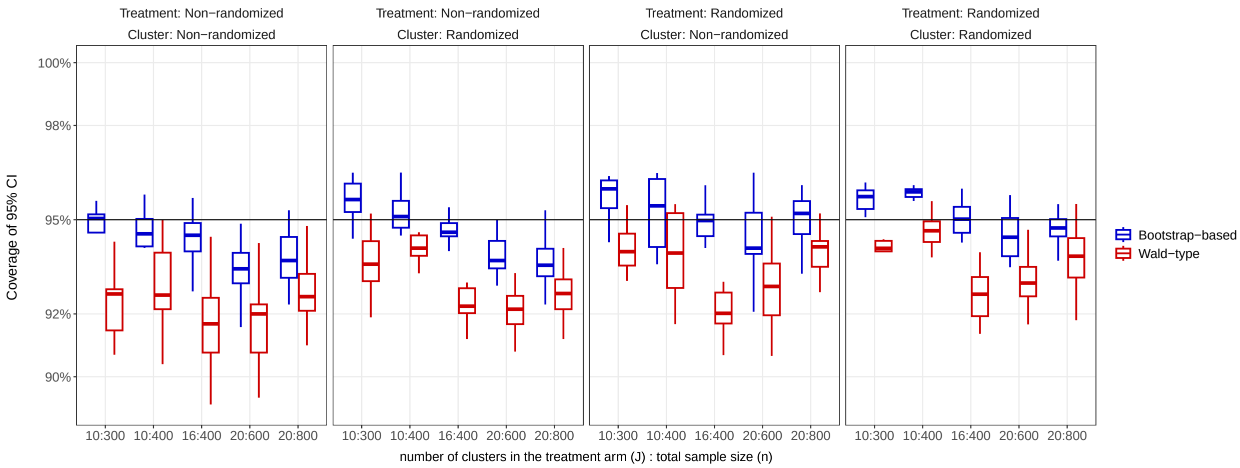
Bias for the Cluster-specific Treatment Effects



Note. |Bias| = Absolute value of the bias.

Figure S2

Coverage Rate for the Cluster-specific Treatment Effects



Note. CI = 95% interval.

Figure S3

Mean Squared Error (MSE) for the Cluster-specific Treatment Effects

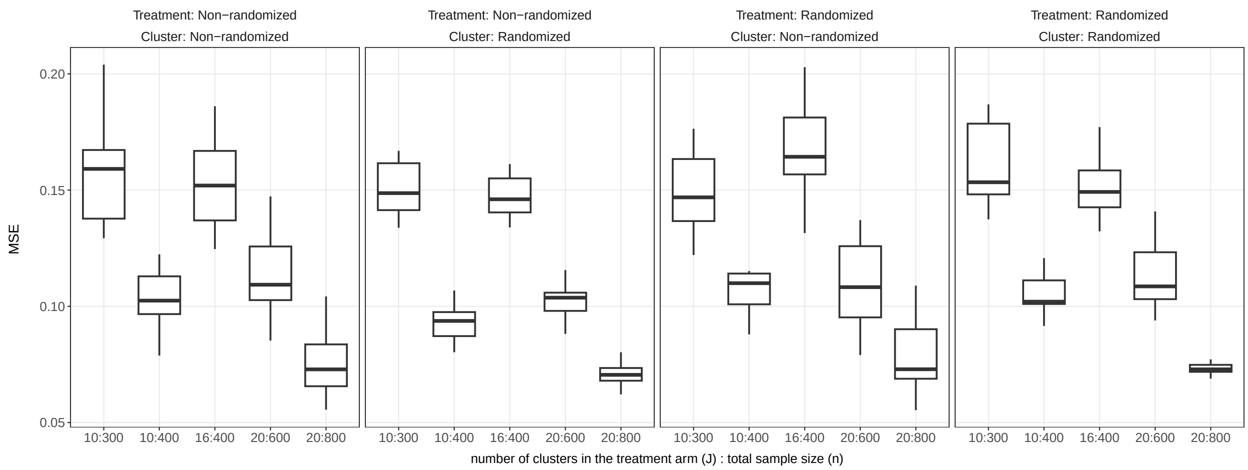
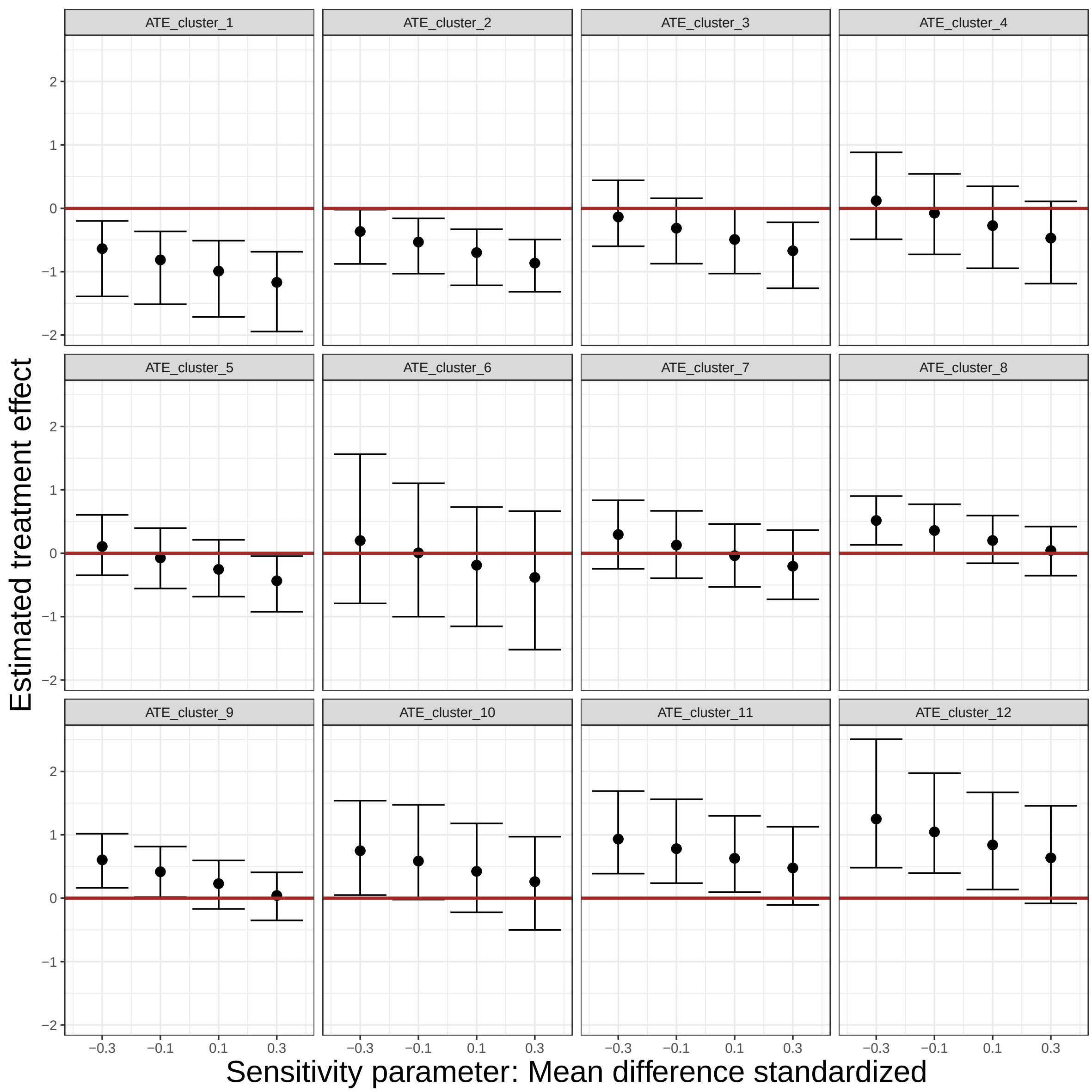


Figure S4

Sensitivity Analysis for the Principal Ignorability Assumption (IA2) in the Empirical Example



Note. For each cluster k ("ATE_cluster_k"), the plots show the estimate and 95% interval of the cluster-specific treatment effect under varying degrees of deviation from the principal ignorability assumption specified by the sensitivity parameter. The sensitivity parameter is $\mathsf{diff}_k = \mathrm{E}[Y(0) \mid K(1) = k, X] - \mathrm{E}[Y(0) \mid K(1) \neq k, X]$, which is the conditional mean difference of the potential control outcome between cluster k and the other clusters. The varying values of the sensitivity parameter are $\mathsf{diff}_k / s_y = -0.3, -0.1, 0.1$, and 0.3, where s_y is the sample standard deviation of the outcome of the control arm.

The robustness of the estimator

For notation simplicity, we omit subscript i for individual i here. In the estimator, we can write the asymptotic bias of the denominator as: $bias^{\text{denominator}} = \mathrm{E}\left[\mathrm{E}\left[K = k \mid T = 1, \mathbf{X}\right] - \hat{p}_k^{\mathrm{dr}}(\mathbf{X}) \mid \mathbf{X}\right]\right]$. The (inner) conditional expectation is further written as $\mathrm{E}\left[p_k(\mathbf{X}) - \hat{p}_k^{\mathrm{dr}}(\mathbf{X}) \mid \mathbf{X}\right] = p_k(\mathbf{X}) - \frac{\pi_t(\mathbf{X})[p_k(\mathbf{X}) - \hat{p}_k(\mathbf{X})]}{\widehat{\pi_t}(\mathbf{X})} - \hat{p}_k(\mathbf{X})$, which (by rearranging terms) equals $\frac{[\pi_t(\mathbf{X}) - \widehat{\pi_t}(\mathbf{X})][p_k(\mathbf{X}) - \hat{p}_k(\mathbf{X})]}{\widehat{\pi_t}(\mathbf{X})}$. Thus, the bias for the denominator is

$$bias^{\text{denominator}} = \mathrm{E}\left[\frac{[\pi_t(\mathbf{X}) - \widehat{\pi_t}(\mathbf{X})][p_k(\mathbf{X}) - \hat{p}_k(\mathbf{X})]}{\widehat{\pi_t}(\mathbf{X})}\right],$$

which is (asymptotically) 0 (i) if $\pi_t(\mathbf{X}) - \widehat{\pi_t}(\mathbf{X}) = 0$, which would be achieved asymptotically when $\widehat{\pi_t}(\mathbf{X})$ is obtained with a correctly specified parametric model for the true treatment probability $\pi_t(\mathbf{X})$, or (ii) if $p_k(\mathbf{X}) - \hat{p}_k(\mathbf{X}) = 0$, which would be achieved asymptotically when $\hat{p}_k(\mathbf{X})$ is obtained with a correctly specified parametric model for the true cluster assignment probability $p_k(\mathbf{X})$, or if (iii) $\pi_t(\mathbf{X}) - \widehat{\pi_t}(\mathbf{X}) = 0$ and $p_k(\mathbf{X}) - \hat{p}_k(\mathbf{X}) = 0$, which would be achieved asymptotically when $\widehat{\pi_t}(\mathbf{X})$ and $\hat{p}_k(\mathbf{X})$ are both obtained with the correctly specified parametric models.

We further express the biases of the numerator of the estimator. Let $\widehat{\phi_1} = \mathrm{E}\left[\frac{1_{\{T=1\}}1_{\{K=k\}}}{\widehat{\pi_t}(\mathbf{X})}\left\{Y - \hat{\mu}_{y|1,k}(\mathbf{X})\right\} + \hat{\mu}_{y|1,k}(\mathbf{X})\hat{p}_k^{\mathrm{dr}}(\mathbf{X})\right]$ and $\widehat{\phi_0} = \mathrm{E}\left[\frac{\hat{p}_k(\mathbf{X})\mathbf{1}_{(T=0)}}{1-\hat{\pi}_t(\mathbf{X})}\left\{Y - \hat{\mu}_{y|0}(\mathbf{X})\right\} + \hat{\mu}_{y|0}(\mathbf{X})\hat{p}_k^{\mathrm{dr}}(\mathbf{X})\right]$. The bias is then a combination of the biases of $\widehat{\phi_1}$ and $\widehat{\phi_0}$. Below, we express the two bias terms to show that either bias term is zero if two or three of the nuisance functions are correct.

For the bias of $\widehat{\phi_1}$, the first term in $\widehat{\phi_1}$ is consistent for $\mathrm{E}\left[\frac{1_{\{T=1\}}1_{\{K=k\}}}{\widehat{\pi_t}(\mathbf{X})}\left\{Y - \hat{\mu}_{y|1,k}(\mathbf{X})\right\}\right] = \mathrm{E}[\mathrm{E}\left[\frac{1_{\{T=1\}}1_{\{K=k\}}}{\widehat{\pi_t}(\mathbf{X})}\left\{Y - \hat{\mu}_{y|1,k}(\mathbf{X})\right\} \mid T = 1, K = 1, \mathbf{X}\right]] = \mathrm{E}\left[\frac{\pi_t(\mathbf{X})p_k(\mathbf{X})}{\widehat{\pi_t}(\mathbf{X})}\left\{\mu_{y|1,k}(\mathbf{X}) - \hat{\mu}_{y|1,k}(\mathbf{X})\right\}\right]$. The second term in $\widehat{\phi_1}$ is consistent for $\mathrm{E}\left[\hat{\mu}_{y|1,k}(\mathbf{X})\hat{p}_k^{\mathrm{dr}}(\mathbf{X})\right] = \mathrm{E}[\mathrm{E}\left[\hat{\mu}_{y|1,k}(\mathbf{X})\left[\frac{1_{\{T=1\}}[1_{\{K=k\}}-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})\right] \mid T=1, K=1, \mathbf{X}\right]] = \mathrm{E}\left[\hat{\mu}_{y|1,k}(\mathbf{X})[\frac{\pi_t(\mathbf{X})[p_k(\mathbf{X})-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})]\right]$.

Thus, the bias term $\widehat{\phi_1} - \mathrm{E}\left[\mu_{y|1,k}(\mathbf{X})p_k(\mathbf{X})\right]$ is consistent for $\mathrm{E}\left[\frac{\pi_t(\mathbf{X})p_k(\mathbf{X})}{\widehat{\pi_t}(\mathbf{X})}\{\mu_{y|1,k}(\mathbf{X}) - \hat{\mu}_{y|1,k}(\mathbf{X})\} + \hat{\mu}_{y|1,k}(\mathbf{X})\left[\frac{\pi_t(\mathbf{X})[p_k(\mathbf{X})-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})\right] - Y(1)1_{\{K(1)=k\}}\right]$. We further express it by conditioning on the covariates, $\mathrm{E}\left[bias^{\phi_1}(\mathbf{X})\right]$, where $bias^{\phi_1}(\mathbf{X}) = \frac{\pi_t(\mathbf{X})p_k(\mathbf{X})}{\widehat{\pi_t}(\mathbf{X})}\{\mu_{y|1,k}(\mathbf{X}) - \hat{\mu}_{y|1,k}(\mathbf{X})\} + \hat{\mu}_{y|1,k}(\mathbf{X})\left[\frac{\pi_t(\mathbf{X})[p_k(\mathbf{X})-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})\right] - \mu_{y|1,k}(\mathbf{X})p_k(\mathbf{X})$; by rearranging the terms, we obtain: $bias^{\phi_1}(\mathbf{X}) = \frac{[\mu_{y|1,k}(\mathbf{X})p_k(\mathbf{X})-\hat{\mu}_{y|1,k}(\mathbf{X})\widehat{p_k}(\mathbf{X})][\pi_t(\mathbf{X})-\hat{\pi}_t(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})}$.

Thus, the bias term $\widehat{\phi_1} - \mathrm{E}\left[\mu_{y|1,k}(\mathbf{X})p_k(\mathbf{X})\right] = \mathrm{E}\left[bias^{\phi_1}(\mathbf{X})\right]$ is asymptotically 0, (i) if the estimator of the treatment probability is correctly specified so that asymptotically $\pi_t(\mathbf{X}) - \hat{\pi}_t(\mathbf{X}) = 0$, or (ii) if the estimators of the cluster assignment probability and outcome mean are correctly specified such that asymptotically $\hat{\mu}_{y|1,k}(\mathbf{X})\widehat{p_k}(\mathbf{X}) - \mu_{y|1,k}(\mathbf{X})p_k(\mathbf{X}) = 0$, or (iii) if the estimators of all of the three nuisance functions are correctly specified.

Similarly, for the bias of $\widehat{\phi_0}$, the first term is consistent for $\mathrm{E}\left[\frac{\hat{p}_k(\mathbf{X})(1-\pi_t(\mathbf{X}))}{1-\hat{\pi}_t(\mathbf{X})}\{\mu_{y|0}(\mathbf{X}) - \hat{\mu}_{y|0}(\mathbf{X})\}\right]$, and the second term is consistent for $\mathrm{E}\left[\hat{\mu}_{y|0}(\mathbf{X})[\frac{\pi_t(\mathbf{X})[p_k(\mathbf{X})-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})]\right]$.

Thus, the bias term $\widehat{\phi_0} - \mathrm{E}\left[\mu_{y|0}(\mathbf{X})p_k(\mathbf{X})\right]$ is expressed as $\mathrm{E}\left[bias^{\phi_0}(\mathbf{X})\right]$, where $bias^{\phi_0}(\mathbf{X}) = \frac{\hat{p}_k(\mathbf{X})(1-\pi_t(\mathbf{X}))}{1-\hat{\pi}_t(\mathbf{X})}\{\mu_{y|0}(\mathbf{X}) - \hat{\mu}_{y|0}(\mathbf{X})\} + \hat{\mu}_{y|0}(\mathbf{X})[\frac{\pi_t(\mathbf{X})[p_k(\mathbf{X})-\widehat{p_k}(\mathbf{X})]}{\hat{\pi}_t(\mathbf{X})} + \widehat{p_k}(\mathbf{X})] - \mu_{y|0}(\mathbf{X})p_k(\mathbf{X})$; by rearranging the terms, we have: $bias^{\phi_0}(\mathbf{X}) = \frac{1}{1-\hat{\pi}_t(\mathbf{X})}\{[\mu_{y|0}(\mathbf{X}) - \hat{\mu}_{y|0}(\mathbf{X})][\widehat{p_k}(\mathbf{X})(1-\pi_t(\mathbf{X})) - p_k(\mathbf{X})(1-\hat{\pi}_t(\mathbf{X}))]\} + \frac{1}{\hat{\pi}_t(\mathbf{X})}\{\hat{\mu}_{y|0}(\mathbf{X})[\pi_t(\mathbf{X}) - \hat{\pi}_t(\mathbf{X})][p_k(\mathbf{X}) - \widehat{p_k}(\mathbf{X})]\}$.

Thus, the bias term $\widehat{\phi_0} - \mathrm{E}\left[\mu_{y|0}(\mathbf{X})p_k(\mathbf{X})\right] = \mathrm{E}\left[bias^{\phi_0}(\mathbf{X})\right]$ is asymptotically 0, (i) if the estimator of the outcome mean and treatment probability are correctly specified such that asymptotically $\mu_{y|0}(\mathbf{X}) - \hat{\mu}_{y|0}(\mathbf{X}) = 0$ and $\pi_t(\mathbf{X}) - \hat{\pi}_t(\mathbf{X}) = 0$, or (ii) if the outcome mean and cluster assignment probability are correctly specified such that asymptotically $\mu_{y|0}(\mathbf{X}) - \hat{\mu}_{y|0}(\mathbf{X}) = 0$ and $p_k(\mathbf{X}) - \widehat{p_k}(\mathbf{X}) = 0$, or (iii) if the treatment probability and cluster assignment probability are are correctly specified such that asymptotically $\widehat{p_k}(\mathbf{X})\left(1 - \pi_t(\mathbf{X})\right) - p_k(\mathbf{X})\left(1 - \hat{\pi}_t(\mathbf{X})\right) = 0$ and $[\pi_t(\mathbf{X}) - \hat{\pi}_t(\mathbf{X})][p_k(\mathbf{X}) - \widehat{p_k}(\mathbf{X})] = 0$, or (iv) if the estimators of all of the three nuisance functions are correctly specified.

References

Ding, P., & Lu, J. (2017). Principal stratification analysis using principal scores. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *79*(3), 757–777. https://doi.org/10.1111/rssb.12191

Jiang, Z., Yang, S., & Ding, P. (2022). Multiply robust estimation of causal effects under principal ignorability. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *84*(4), 1423–1445. https://doi.org/10.1111/rssb.12538

Kennedy, E. H. (2023, January). Semiparametric doubly robust targeted double machine learning: A review [arXiv:2203.06469 [stat]]. Retrieved November 30, 2023, from http://arxiv.org/abs/2203.06469

Nguyen, T. Q., Stuart, E. A., Scharfstein, D. O., & Ogburn, E. L. (2023, October). Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects [arXiv:2303.05032 [stat]]. https://doi.org/10.48550/arXiv.2303.05032