Supplementary Material for Imai, Jiang, and Malani. "Causal Inference with Interference and Noncompliance in Two-Stage Randomized Experiments."

The supplementary material contains the following five sections:

- Section A provides the results for the complier average spillover effects under stratified interference.
- Section B gives the proofs for the randomization-based inference approach.
- Section C gives the proofs for the regression-based approach.
- Section D presents the simulation studies.
- Section E proposes a model-based approach to overcome the limitations of the nonparametric approach in the main text.

A Complier Average Spillover Effects under Stratified Interference

Stratified interference, i.e., Assumption 6, allows us to define the complier average spillover effect (CASE), representing the average causal effect of treatment assignment mechanism among compliers while holding their own treatment assignment at a fixed value,

$$\mathsf{CASE}(z) = \frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(z,1)-Y_{ij}(z,0)\}I\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}I\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\}}.$$

We emphasize that this estimand is defined only when the spillover effect of treatment assignment on the treatment receipt is present (otherwise, the denominator is zero). Note that the compliers here are defined differently than those for the CADE. Specifically, the compliers for the CASE are those who receive the treatment only when the assignment mechanism A_j is equal to 1, i.e., units with $\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\}$. Thus, the CASE represents the average causal effect of the assignment mechanism on the outcome among the compliers while holding their treatment assignment status constant.

A.1 Nonparametric Identification

To establish the nonparametric identification result for the CASE, we need two assumptions similar to Assumptions 4 and 5 for the CADE.

Assumption A1 (Monotonicity with Respect to the Assignment Mechanism)

$$D_{ij}(z,1) \geq D_{ij}(z,0) \quad \text{for all } z=0,1.$$

The assumption states that a unit is no less likely to receive the treatment under the treatment assignment mechanism $A_j=1$ than under the treatment assignment mechanism

$A_j = 0$, holding its own treatment assignment at a constant. In our application, Assumption A1 implies that a household is no less likely to enroll in the RSBY when a greater number of households are encouraged to do so.

Next, we introduce the assumption of restricted interference similar to Assumption 5.

ASSUMPTION A2 (RESTRICTED INTERFERENCE UNDER NONCOMPLIANCE FOR THE ASSIGNMENT MECHANISM) *Under Assumption 6, for a given unit i in cluster j, if $D_{ij}(z,1) = D_{ij}(z,0)$ for some z, then $Y_{ij}(\boldsymbol{D}_j(z,1)) = Y_{ij}(\boldsymbol{D}_j(z,0))$.*

The assumption states that if the treatment receipt of a unit is not affected by the assignment mechanism of the cluster, its outcome should also not be affected by the assignment mechanism. Similar to Assumption 5, this assumption holds in case of no spillover effect of treatment receipt on the outcome (equation (3)). As noted above, however, in case of no spillover effect on the treatment receipt (equation (4)), the CASE is not well defined. Furthermore, when both spillover effects are present, Assumption A2 is likely to be violated.

We provide the nonparametric identification and consistent estimation results for the CASE that are analogous to those presented in Theorem 3 for the CADE.

THEOREM A1 (NONPARAMETRIC IDENTIFICATION AND CONSISTENT ESTIMATION OF THE COMPLIER AVERAGE SPILLOVER EFFECT UNDER STRATIFIED INTERFERENCE) *Suppose that the outcome is bounded. Then, under Assumptions 1–3, 6 and A1–A2, we have*

$$\lim_{n_j \to \infty, J \to \infty} \mathsf{CASE}(z) \;=\; \operatorname*{plim}_{n_j \to \infty, J \to \infty} \frac{\widehat{\mathsf{SEY}}(z)}{\widehat{\mathsf{SED}}(z)}$$

for $z = 0, 1$.

Proof is in Appendix B.4.

A.2 Effect Decomposition

We consider the following decomposition of the CASE analogous to that of the CADE,

$$\mathsf{SEY}(z) \;=\; \mathsf{CASE}(z) \cdot \lambda_c(z) + \mathsf{NASE}(z) \cdot \{1 - \lambda_c(z)\}, \tag{A1}$$

where the average noncomplier spillover effect is defined as,

$$\mathsf{NASE}(z) \;=\; \frac{\sum_{j=1}^{J} \sum_{i=1}^{n_j} \{Y_{ij}(z,1) - Y_{ij}(z,0)\} I\{D_{ij}(z,1) = D_{ij}(z,0)\}}{\sum_{j=1}^{J} \sum_{i=1}^{n_j} I\{D_{ij}(1,a) = D_{ij}(0,a)\}},$$

and the proportion of compliers with respect to the treatment assignment is given by,

$$\lambda_c(z) \;=\; \frac{1}{N} \sum_{j=1}^{J} \sum_{i=1}^{n_j} I\{D_{ij}(z,1) = 1, D_{ij}(z,0) = 0\}.$$

For compliers with $D_{ij}(z_{ij},1) = 1$ and $D_{ij}(z_{ij},0) = 0$, the exclusion restriction, i.e., Assumption 3, implies the following decomposition,

$$\begin{aligned} Y_{ij}(z_{ij},1) - Y_{ij}(z_{ij},0) \;=\; & \{Y_{ij}(D_{ij} = 1, \boldsymbol{D}_{-i,j}(z_{ij},1)) - Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j}(z_{ij},0))\} \\ & + \{Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j}(z_{ij},1)) - Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j}(z_{ij},0))\}, \end{aligned}$$

which shows that the effect of treatment assignment mechanism on the outcome for a unit equals the sum of the direct effect through its own treatment receipt and the indirect effect through the treatment receipts of the other units in the same cluster. For noncompliers with $D_{ij}(z_{ij}, 1) = D_{ij}(z_{ij}, 0) = d$ where $d = 0, 1$, we can write the total effect of treatment assignment mechanism as,

$$\begin{aligned} & Y_{ij}(z_{ij}, 1) - Y_{ij}(z_{ij}, 0) \\ = \quad & Y_{ij}(D_{ij} = d, \boldsymbol{D}_{-i,j}(Z_{ij} = z_{ij}, 1)) - Y_{ij}(D_{ij} = d, \boldsymbol{D}_{-i,j}(Z_{ij} = z_{ij}, 0)), \end{aligned}$$

which characterizes the spillover effect of the treatment assignments of the other units on the outcome through their treatment receipts. Assumption A2 guarantees this effect is zero for noncompliers, implying $\mathsf{NASE}(z) = 0$ and the identification of $\mathsf{CASE}(z)$.

A.3 Randomization-based Variances

We can also derive the randomization-based variances of the proposed spillover effect estimators. We begin by defining the following quantities,

$$\begin{aligned} \sigma_b^2(z, a) &= \frac{1}{J-1} \sum_{j=1}^{J} \left\{ \frac{n_j J}{N} \overline{Y}_j(z, a) - \overline{Y}(z, a) \right\}^2, \\ \sigma_{SE}^2(z) &= \frac{1}{J-1} \sum_{j=1}^{J} \left\{ \frac{n_j J}{N} \mathsf{SEY}_j(z) - \mathsf{SEY}(z) \right\}^2, \end{aligned}$$

where $\sigma_b^2(z, a)$ is the between-cluster variance of $Y_{ij}(z, a)$, and $\sigma_{SE}^2(z)$ is the between-cluster variance of $\mathsf{SEY}_{ij}(a)$. The next theorem presents the randomization-based variance.

THEOREM A2 (VARIANCES OF THE ITT SPILLOVER EFFECT ESTIMATORS) *Under Assumptions 1, 2 and 6, we have*

$$\text{var}\left\{\widehat{\mathsf{SEY}}(z)\right\} = \frac{\sigma_b^2(z,1)}{J_1} + \frac{\sigma_b^2(z,0)}{J_0} - \frac{\sigma_{SE}^2(z)}{J} + \sum_{a=0}^{1} \frac{1}{J_a J} \sum_{j=1}^{J} \text{var}\left\{\widehat{Y}_j(z,a) \mid A_j = a\right\},$$

where

$$\text{var}\left\{\widehat{Y}_j(z,a) \mid A_j = a\right\} = \frac{1}{n_{jz}} \left(1 - \frac{n_{jz}}{n_j}\right) \sigma_j^2(z,a).$$

In contrast to the case of the direct effect estimators, the variances of the spillover effect estimators are based on cluster-robust variances alone. However, no unbiased estimation of the variance of $\widehat{\mathsf{SEY}}(z)$ is available because $\sigma_{SE}^2(z)$ is not identifiable. Hence, we propose a conservative estimator of the variance,

$$\widehat{\text{var}}\left\{\widehat{\mathsf{SEY}}(z)\right\} = \frac{\widehat{\sigma}_b^2(z,1)}{J_1} + \frac{\widehat{\sigma}_b^2(z,0)}{J_0},$$

which is no less than the true variance in expectation, i.e., $\mathbb{E}\left[\widehat{\text{var}}\left\{\widehat{\mathsf{SEY}}(z)\right\}\right] \geq \text{var}\left\{\widehat{\mathsf{SEY}}(z)\right\}$. The inequality becomes equality when the cluster-level spillover effect, i.e., $n_j J\{\overline{Y}_j(z,1) - \overline{Y}_j(z,0)\}/N$, is constant (see Appendix B.7 for a proof).

Finally, the variance of the CASE estimators can be derived by applying the Delta method as done in Theorem 5. The resulting variances involve the covariance between $\widehat{\mathsf{SEY}}$ and $\widehat{\mathsf{SED}}$ whose expression is shown in Appendix B.6.

B Proofs for the Randomization-Based Inference Approach

B.1 Testable Conditions for No Spillover Effect of Treatment Receipt on the Outcome

When there is no spillover effect of treatment receipt on the outcome, we define

$$
\begin{aligned}
& \overline{(Y=y)D}_{ij}(z,a) \\
= & \sum_{\boldsymbol{z}_{-i,j}\in\mathcal{Z}_{-i,j}} I\{Y_{ij}(Z_{ij}=z,\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j})=y\}D_{ij}(Z_{ij}=z,\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j}) \\
& \qquad \cdot \Pr(\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j}\mid Z_{ij}=z,A_j=a) \\
= & \sum_{\boldsymbol{z}_{-i,j}\in\mathcal{Z}_{-i,j}} I\{Y_{ij}(D_{ij}=1)=y\}D_{ij}(Z_{ij}=z,\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j})\Pr(\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j}\mid Z_{ij}=z,A_j=a)
\end{aligned}
$$

for any y. Because

$$
\lim_{n_j\to\infty}\{\Pr(\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j}\mid Z_{ij}=1,A_j=a)-\Pr(\boldsymbol{Z}_{-i,j}=\boldsymbol{z}_{-i,j}\mid Z_{ij}=0,A_j=a)\} \;=\; 0,
$$

we can obtain $\lim_{n_j\to\infty}\overline{(Y=y)D}_{ij}(1,a)\ \geq\ \lim_{n_j\to\infty}\overline{(Y=y)D}_{ij}(0,a)$ under Assumption 4. As a result, we have $\overline{(Y=y)D}(1,a)\geq\overline{(Y=y)D}(0,a)$, where

$$
\overline{(Y=y)D}(z)=\frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\overline{(Y=y)D}_{ij}(z).
$$

Similarly, we can obtain $\lim_{n_j\to\infty}\overline{(Y=y)(1-D)}(1,a)\geq\lim_{n_j\to\infty}\overline{(Y=y)(1-D)}(0,a)$. Therefore, we obtain the following testable conditions for no spillover effect of the treatment receipt on the outcome,

$$
\begin{aligned}
\lim_{n_j\to\infty}\overline{(Y=y)D}_{ij}(1,a) &\geq \lim_{n_j\to\infty}\overline{(Y=y)D}_{ij}(0,a), \\
\lim_{n_j\to\infty}\overline{(Y=y)(1-D)}(1,a) &\geq \lim_{n_j\to\infty}\overline{(Y=y)(1-D)}(0,a).
\end{aligned}
$$

Similar to the unbiased estimation of the ITT effects, we can use

$$
\begin{aligned}
\widehat{(Y=y)D}(z,a) &= \frac{\frac{1}{N}\sum_{j=1}^{J}n_j\widehat{(Y=y)D}_j(z,a)I(A_j=a)}{\frac{1}{J}\sum_{j=1}^{J}I(A_j=a)}, \\
\widehat{(Y=y)(1-D)}(z,a) &= \frac{\frac{1}{N}\sum_{j=1}^{J}n_j\widehat{(Y=y)(1-D)}_j(z,a)I(A_j=a)}{\frac{1}{J}\sum_{j=1}^{J}I(A_j=a)},
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{(Y=y)D}_j(z,a) &= \frac{\sum_{i=1}^{n_j}I(Y_{ij}=y)D_{ij}I(Z_{ij}=z)}{\sum_{i=1}^{n_j}I(Z_{ij}=z)}, \\
\widehat{(Y=y)(1-D)}_j(z,a) &= \frac{\sum_{i=1}^{n_j}I(Y_{ij}=y)(1-D_{ij})I(Z_{ij}=z)}{\sum_{i=1}^{n_j}I(Z_{ij}=z)}
\end{aligned}
$$

to unbiasedly estimate $\overline{(Y=y)D}(z,a)$ and $\overline{(Y=y)(1-D)}(z,a)$, respectively. As a result, we can use the observed data to test whether there is a spillover effect of treatment receipt on the outcome.

B.2 Proof of Theorem 2

We first prove the nonparametric identification as the cluster size n_j goes to infinity for each j. Under this scenario, the treatment assignment of one unit becomes asymptotically independent of another unit's treatment assignment given its treatment assignment mechanism within the same cluster. This yields

$$\lim_{n_j \to \infty} \{\Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid Z_{ij} = z, A_j = a) - \Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid A_j = a)\} \ = \ 0$$

for $z = 0, 1$. Therefore, we have,

$$\begin{aligned} &\lim_{n_j \to \infty} \mathsf{DED}(a) \\ = \ &\lim_{n_j \to \infty} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \sum_{\boldsymbol{z}_{-i,j} \in \mathcal{Z}_{-i,j}} \{D_{ij}(1, \boldsymbol{z}_{-i,j}) - D_{ij}(0, \boldsymbol{z}_{-i,j})\} \Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid A_j = a) \\ = \ &\lim_{n_j \to \infty} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \sum_{\boldsymbol{z}_{-i,j} \in \mathcal{Z}_{-i,j}} C_{ij}(\boldsymbol{z}_{-i,j}) \Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid A_j = a), \end{aligned} \tag{A2}$$

where the last equality follows from Assumption 4. Next, we show that the numerator of $\mathsf{CADE}(a)$ is equal to,

$$\begin{aligned} &\lim_{n_j \to \infty} \mathsf{DEY}(a) \\ = \ &\lim_{n_j \to \infty} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \sum_{\boldsymbol{z}_{-i,j} \in \mathcal{Z}_{-i,j}} \{Y_{ij}(1, \boldsymbol{z}_{-i,j}) - Y_{ij}(0, \boldsymbol{z}_{-i,j})\} \Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid A_j = a) \\ = \ &\lim_{n_j \to \infty} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \sum_{\boldsymbol{z}_{-i,j} \in \mathcal{Z}_{-i,j}} \{Y_{ij}(1, \boldsymbol{z}_{-i,j}) - Y_{ij}(0, \boldsymbol{z}_{-i,j})\} C_{ij}(\boldsymbol{z}_{-i,j}) \Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid A_j = a), \end{aligned} \tag{A3}$$

where the last equality follows from Assumption 5. Thus, we obtain the desired result,

$$\lim_{n_j \to \infty} \mathsf{CADE}(a) \ = \ \lim_{n_j \to \infty} \frac{\mathsf{DEY}(a)}{\mathsf{DED}(a)}.$$

Next, we establish the consistent estimation. We assume the following restriction on interference in (Sävje *et al.*, 2017) hold, which still allows the total amount of interference within each cluster to grow with the cluster size,

$$\frac{1}{n_j} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \iota_{ii'} \ = \ o(n_j), \quad \text{where} \quad \iota_{ii'} \ = \ \begin{cases} 1 & \text{if } \mathcal{I}_{\ell i} \mathcal{I}_{\ell i'} = 1 \text{ for some } \ell = 1, 2, \ldots, n_j, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mathcal{I}_{\ell i} \ = \ \begin{cases} 1 & \text{if } D_{ij}(\boldsymbol{z}_j) \neq D_{ij}(\boldsymbol{z}'_j) \quad \text{for } z_{\ell j} \neq z'_{\ell j} \text{ and } \boldsymbol{z}_{-\ell,j} = \boldsymbol{z}'_{-\ell,j}, \\ 1 & \text{if } i = \ell, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\frac{1}{n_j}\sum_{i=1}^{n_j}\sum_{i'=1}^{n_j}\kappa_{ii'} \ = \ o(n_j), \quad \text{where} \quad \kappa_{ii'} \ = \ \begin{cases} 1 & \text{if } K_{\ell i}K_{\ell i'} \text{ for some } \ell = 1,2,\ldots,n_j, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$K_{i\ell} \ = \ \begin{cases} 1 & \text{if } Y_{ij}(\boldsymbol{z}_j) \neq Y_{ij}(\boldsymbol{z}'_j) \quad \text{for } z_{\ell j} \neq z'_{\ell j} \text{ and } \boldsymbol{z}_{-i,j} = \boldsymbol{z}'_{-i,j}, \\ 1 & \text{if } i = \ell, \\ 0 & \text{otherwise.} \end{cases}$$

From Sävje *et al.* (2017), under the proposed conditions, we can consistently estimate the DED and the DEY within each cluster,

$$\operatorname*{plim}_{n_j\to\infty} \widehat{\mathsf{DED}}_j(a) \ = \ \mathsf{DED}_j(a), \quad \operatorname*{plim}_{n_j\to\infty} \widehat{\mathsf{DEY}}_j(a) \ = \ \mathsf{DEY}_j(a) \tag{A4}$$

for all j. Furthermore, because A_j is the sampling indicator of a simple random sampling from $(n_1 J\overline{D}_1(z,a)/N, \ldots, n_J J\overline{D}_J(z,a)/N)$, as the number of clusters also tends to infinity, we have,

$$\operatorname*{plim}_{n_j\to\infty, J\to\infty} \widehat{\mathsf{DED}}(a) \ = \ \mathsf{DED}(a).$$

Similarly, we can obtain

$$\operatorname*{plim}_{n_j\to\infty, J\to\infty} \widehat{\mathsf{DEY}}(a) \ = \ \mathsf{DEY}(a).$$

Putting all together establishes the consistent estimation of the CADE. □

B.3 Asymptotic Normality of the ITT Effect Estimators Under Stratified Interference

We provide the conditions for the asymptotic normality of the estimator of the ITT effect on the outcome. The conditions for the estimator of the ITT effect on the treatment receipt can be obtained in a similar fashion. Ohlsson (1989) establishes the asymptotic normality for two-stage sampling from a finite population. In our setting, we generalize his result to the two-stage randomized experiments by verifying the conditions required by his result in our context. We first state the following finite population central limit theorem.

THEOREM A3 (FINITE POPULATION CENTRAL LIMIT THEOREM (HÁJEK, 1960)) *Let $\bar{v}_S$ be the average of a simple random sample of size n from a finite population $\{v_1, \ldots, v_N\}$. As $N \to \infty$, if*

$$\frac{1}{\min(n, N-n)} \cdot \frac{\max_{1\le i\le N}(v_i - \bar{v}_N)^2}{\sum_{i=1}^{N}(v_i - \bar{v}_N)^2/(N-1)} \longrightarrow 0, \tag{A5}$$

where $\bar{v}_N$ is the average of the population, then $(\bar{v}_S - \bar{v}_N)/\sqrt{var(\bar{v}_S)} \xrightarrow{d} N(0,1)$.

Equation (A5) holds if v_i's are bounded and n and N go to infinity.

We next introduce the central limit theorem under two-stage sampling.

THEOREM A4 (FINITE POPULATION CENTRAL LIMIT THEOREM UNDER TWO-STAGE SAMPLING (OHLSSON, 1989)) *Let v_{ij} be the outcome of interest of unit i in cluster j, where $i = 1, \ldots, n_j$ and $J = 1, \ldots, J$. Define $v_{\cdot j} = \sum_{i=1}^{n_j} v_{ij}$. Clusters are sampled from the population in the first stage and units are sampled in the second stage within the sampled clusters from the first stage. Let W_j be the sample indicator of the first stage and I_{ij} be the sample indicator of the second stage. Define $T_j = \sum_{i=1}^{J} I_{ij} v_{ij}$, $T = \sum_{j=1}^{J} W_j T_j$, $Q = \sum_{j=1}^{J} W_j v_{\cdot j}$, $\sigma_j^2 = \text{var}(T_i \mid W_1, \ldots, W_J)$ and $\mu_j^{(4)} = \mathbb{E}\{(T_j - v_{\cdot j})^4 \mid W_1, \ldots, W_J\}$. Then,*

$$\frac{T - \mathbb{E}(T)}{\sqrt{\text{var}(T)}} \xrightarrow{d} N(0, 1)$$

if the following three conditions hold

$$\frac{Q - \mathbb{E}(Q)}{\sqrt{\text{var}(Q)}} \quad \xrightarrow{d} \quad N(0, 1), \tag{A6}$$

$$\frac{\sum_{j=1}^{J} \mu_j^{(4)} \mathbb{E}(W_j^4)}{\sum_{j=1}^{J} \sigma_j^2 \mathbb{E}(W_j^2)} \quad \longrightarrow \quad 0, \tag{A7}$$

$$\text{cov}(W_j^2, W_{j'}^2) \quad \leq \quad 0, \text{ for } j \neq j'. \tag{A8}$$

To apply Theorem A4, we decompose $\widehat{\mathsf{DEY}}(a)$ as,

$$\widehat{\mathsf{DEY}}(a) = \frac{1}{J_a} \sum_{j=1}^{J} I(A_j = a) \cdot \frac{n_j J}{N} \mathsf{DEY}_j(a) + \frac{1}{J} \sum_{j=1}^{J} I(A_j = a) \cdot \frac{n_j J}{N} \{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}, \tag{A9}$$

and $\widehat{\mathsf{SEY}}(z)$ as,

$$\begin{aligned} \widehat{\mathsf{SEY}}(z) = {} & \sum_{j=1}^{J} A_j \cdot \frac{n_j J}{N} \left\{ \frac{\overline{Y}_j(z,1)}{J_1} + \frac{\overline{Y}_j(z,0)}{J_0} \right\} - \sum_{j=1}^{J} \frac{\overline{Y}_j(z,0)}{J_0} \\ & + \frac{1}{J_1} \sum_{j=1}^{J} A_j \cdot \frac{n_j J}{N} \{\widehat{Y}_j(z,1) - \overline{Y}_j(z,1)\} - \frac{1}{J_0} \sum_{j=1}^{J} (1 - A_j) \cdot \frac{n_j J}{N} \{\widehat{Y}_j(z,0) - \overline{Y}_j(z,0)\}. \end{aligned} \tag{A10}$$

Denote the first part of each equation above as,

$$\widehat{\mathsf{DEY}}^{\text{cluster}}(a) = \frac{1}{J_a} \sum_{j=1}^{J} I(A_j = a) \cdot \frac{n_j J}{N} \mathsf{DEY}_j(a),$$

$$\widehat{\mathsf{SEY}}^{\text{cluster}}(z) = \sum_{j=1}^{J} A_j \cdot \frac{n_j J}{N} \left\{ \frac{\overline{Y}_j(z,1)}{J_1} + \frac{\overline{Y}_j(z,0)}{J_0} \right\} - \sum_{j=1}^{J} \frac{\overline{Y}_j(z,0)}{J_0}.$$

We can treat $\widehat{\mathsf{DEY}}^{\text{cluster}}(a)$ and $\widehat{\mathsf{SEY}}^{\text{cluster}}(z)$ as Q in Theorem A4. For $\widehat{\mathsf{DEY}}(a)$, we treat $I(A_j = a)$ as W_j and $\frac{n_j J}{N J_a} \mathsf{DEY}_{ij}(a)$ as v_{ij} in Theorem A4; for $\widehat{\mathsf{SEY}}(z)$, we treat A_j as W_j and $\frac{n_j J}{N} \left\{ \frac{Y_{ij}(z,1)}{J_1} + \frac{Y_{ij}(z,0)}{J_0} \right\}$ as v_{ij} in Theorem A4. We first give the regularity conditions for the asymptotic normality of $\widehat{\mathsf{DEY}}^{\text{cluster}}(a)$ and $\widehat{\mathsf{SEY}}^{\text{cluster}}(z)$:

(a) Equation (A5) holds for $n = J_1$, $N = J$ and $v_i = \frac{n_j J}{N}\overline{Y}_j(z,a)$ for $z = 0,1$ and $a = 0,1$.

(b) Equation (A5) holds for $n = J_1$, $N = J$ and $v_i = \frac{n_j J}{N}\overline{Y}_j(z,1)/J_1 + \frac{n_j J}{N}\overline{Y}_j(z,0)/J_0$ for $z = 0,1$.

For a bounded outcome, these two conditions are satisfied as the number of clusters goes to infinity. According to Theorem A3, under Condition (a), as $J \to \infty$, we have

$$\frac{\widehat{\mathsf{DEY}}^{\text{cluster}}(a) - \mathsf{DEY}(a)}{\sqrt{\text{var}\{\widehat{\mathsf{DEY}}^{\text{cluster}}(a)\}}} \xrightarrow{d} N(0,1), \tag{A11}$$

and under Condition (b), as $J \to \infty$, we have,

$$\frac{\widehat{\mathsf{SEY}}^{\text{cluster}}(z) - \mathsf{SEY}(z)}{\sqrt{\text{var}\{\widehat{\mathsf{SEY}}^{\text{cluster}}(z)\}}} \xrightarrow{d} N(0,1). \tag{A12}$$

Second, we require the following conditions,

(c) As $J \to \infty$

$$\frac{\sum_{j=1}^{J} \mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^4 \mid A_j = a]\text{pr}(A_j = a)}{\left(\sum_{j=1}^{J} \mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^2 \mid A_j = a]\text{pr}(A_j = a)\right)^2} \longrightarrow 0$$

for $a = 0,1$.

(d) As $J \to \infty$

$$\frac{\sum_{j=1}^{J} \mathbb{E}[\{\widehat{Y}_j(z,a) - \overline{Y}_j(z,a)\}^4 \mid A_j = a]\text{pr}(A_j = a)}{\left(\sum_{j=1}^{J} \mathbb{E}[\{\widehat{Y}_j(z,a)) - \overline{Y}_j(z,a)\}^2 \mid A_j = a]\text{pr}(A_j = a)\right)^2} \longrightarrow 0$$

for $z = 0,1$ and $a = 0,1$.

To give some intuition on when these two conditions hold, we show that Conditions (c) and (d) hold if the outcome is bounded and $\mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^2 \mid A_j = a]$ and $\mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^2 \mid A_j = a]$ are equal across different clusters. In this case, the term on the left hand side of Condition (c) is of the same order as

$$\frac{1}{J} \cdot \frac{\mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^4 \mid A_j = a]}{\mathbb{E}[\{\widehat{\mathsf{DEY}}_j(a) - \mathsf{DEY}_j(a)\}^2 \mid A_j = a]^2\text{pr}(A_j = a)},$$

which converges to zero if the outcome is bounded and J goes to infinity. Therefore, Condition (c) holds. Similar argument applies to Condition (d).

We now verify the conditions in Theorem A4. For $\mathsf{DEY}(a)$, Condition (A6) follows from equation (A11), Condition (A7) follows from Condition (c) above, and Condition (A8) follows from $\text{cov}(A_j, A_{j'}) \le 0$ for $j \ne j'$. Therefore, as $J \to \infty$, we have,

$$\frac{\widehat{\mathsf{DEY}}(a) - \mathsf{DEY}(a)}{\sqrt{\text{var}\left\{\widehat{\mathsf{DEY}}(a)\right\}}} \xrightarrow{d} N(0,1).$$

Similarly, for $\mathsf{SEY}(z)$, Condition (A6) in follows from equation (A12), Condition (A7) follows from Condition (d) above, and Condition (A8) follows from $\text{cov}(A_j, A_{j'}) \le 0$ for $j \neq j'$. Therefore, as $J \to \infty$, we have,

$$\frac{\widehat{\mathsf{SEY}}(z) - \mathsf{DEY}(z)}{\sqrt{\text{var}\left\{\widehat{\mathsf{SEY}}(z)\right\}}} \xrightarrow{d} N(0,1).$$

Analogous conditions can be provided for the the asymptotic normality of $\widehat{\mathsf{DED}}(a)$ and $\widehat{\mathsf{SED}}(z)$. In a similar way, we can show the asymptotic normality for any linear combinations of $\{\widehat{\mathsf{DED}}(a), \widehat{\mathsf{DEY}}(a)\}$ and $\{\widehat{\mathsf{SED}}(z), \widehat{\mathsf{SEY}}(z)\}$. The conditions are satisfied if Y is bounded. As a result, we can further show that

$$\begin{pmatrix} \text{var}\left\{\widehat{\mathsf{DED}}(a)\right\} & \text{cov}\left\{\widehat{\mathsf{DED}}(a), \widehat{\mathsf{DEY}}(a)\right\} \\ \text{cov}\left\{\widehat{\mathsf{DED}}(a), \widehat{\mathsf{DEY}}(a)\right\} & \text{var}\left\{\widehat{\mathsf{DEY}}(a)\right\} \end{pmatrix}^{-1/2} \begin{pmatrix} \widehat{\mathsf{DED}}(a) - \mathsf{DED}(a) \\ \widehat{\mathsf{DEY}}(a) - \mathsf{DEY}(a) \end{pmatrix} \xrightarrow{d} \boldsymbol{N}_2(\boldsymbol{0}_2, \boldsymbol{I}_2),$$

$$\begin{pmatrix} \text{var}\left\{\widehat{\mathsf{SED}}(z)\right\} & \text{cov}\left\{\widehat{\mathsf{SED}}(z), \widehat{\mathsf{SEY}}(z)\right\} \\ \text{cov}\left\{\widehat{\mathsf{SED}}(z), \widehat{\mathsf{SEY}}(z)\right\} & \text{var}\left\{\widehat{\mathsf{SEY}}(z)\right\} \end{pmatrix}^{-1/2} \begin{pmatrix} \widehat{\mathsf{SED}}(z) - \mathsf{SED}(z) \\ \widehat{\mathsf{SEY}}(z) - \mathsf{SEY}(z) \end{pmatrix} \xrightarrow{d} \boldsymbol{N}_2(\boldsymbol{0}_2, \boldsymbol{I}_2).$$

In general, the asymptotic normality of the ITT effects only requires some mild conditions as long as the outcome is bounded and J goes to infinity. We leave the development of more refined CLTs under the two-stage randomized experiments to future work.

Finally, because $\widehat{\mathsf{CADE}}(a)$ equals the ratio of $\widehat{\mathsf{DEY}}(a)$ and $\widehat{\mathsf{DED}}(a)$, and $\widehat{\mathsf{CASE}}(z)$ equals the ratio of $\widehat{\mathsf{SEY}}(a)$ and $\widehat{\mathsf{SED}}(a)$, we can obtain the CLT for the CADE and CASE by applying the Delta method.

B.4 Proof of Theorems 3 and A1

According to the asymptotic normality results shown in Appendix B.3, we have

$$\operatorname*{plim}_{n_j\to\infty, J\to\infty} \widehat{\mathsf{DEY}}(a) = \mathsf{DEY}(a), \qquad \operatorname*{plim}_{n_j\to\infty, J\to\infty} \widehat{\mathsf{DED}}(a) = \mathsf{DED}(a).$$

As a result,

$$\lim_{n_j\to\infty, J\to\infty} \mathsf{CADE}(a) \;=\; \operatorname*{plim}_{n_j\to\infty, J\to\infty} \frac{\widehat{\mathsf{DEY}}(a)}{\widehat{\mathsf{DED}}(a)}.$$

For the CASE, under Assumption A1, we have

$$\mathsf{SED}(z) \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j}\{D_{ij}(z,1) - D_{ij}(z,0)\} \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j} C_{ij}(z).$$

We can then obtain

$$\mathsf{SEY}(z) \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(z,1) - Y_{ij}(z,0)\} \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(z,1) - Y_{ij}(z,1)\}\{D_{ij}(z,1) - D_{ij}(z,0)\}$$

$$= \sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(z,1)-Y_{ij}(z,1)\}I\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\},$$

where the second equality follows from Assumption A2. Thus, we have the following equality,

$$\mathsf{CASE}(z) = \frac{\mathsf{SEY}(z)}{\mathsf{SED}(z)}.$$

Based on the asymptotic normality results shown in Appendix B.3, we have

$$\operatorname*{plim}_{n_j\to\infty,J\to\infty} \widehat{\mathsf{SEY}}(z)=\mathsf{SEY}(z), \qquad \operatorname*{plim}_{n_j\to\infty,J\to\infty} \widehat{\mathsf{SED}}(a)=\mathsf{SED}(a).$$

This implies the nonparametric identification and consistent estimation of the CASE. □

B.5 Proof of Theorem 4

We prove a general version of Theorem 4 using the general weight w_j^* whereas in the main text, we consider the special case with $w_j^* = n_j J/N$. Using this general weight, we can rewrite the causal quantities as follows,

$$\mathsf{DED}(a) = \frac{1}{J}\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\overline{D}_{ij}(1,a)-\overline{D}_{ij}(0,a), \quad \mathsf{SED}(z) = \frac{1}{J}\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\overline{D}_{ij}(z,1)-\overline{D}_{ij}(z,0),$$

$$\mathsf{DEY}(a) = \frac{1}{J}\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\overline{Y}_{ij}(1,a)-\overline{Y}_{ij}(0,a), \quad \mathsf{SEY}(z) = \frac{1}{J}\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\overline{Y}_{ij}(z,1)-\overline{Y}_{ij}(z,0),$$

$$\mathsf{CADE}(a) = \frac{\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\{Y_{ij}(1,a)-Y_{ij}(0,a)\}I\{D_{ij}(1,a)=1,D_{ij}(0,a)=0\}}{\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}I\{D_{ij}(1,a)=1,D_{ij}(0,a)=0\}},$$

$$\mathsf{CASE}(z) = \frac{\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}\{Y_{ij}(z,1)-Y_{ij}(z,1)\}I\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\}}{\sum_{j=1}^{J}\frac{w_j^*}{n_j}\sum_{i=1}^{n_j}I\{D_{ij}(z,1)=1,D_{ij}(z,0)=0\}}.$$

The corresponding estimators for the ITT effects can be written as,

$$\widehat{\mathsf{DED}}(a) = \frac{1}{J_a}\sum_{j=1}^{J}w_j^*\widehat{D}_j(1,a)I(A_j=a)-\frac{1}{J_a}\sum_{j=1}^{J}w_j^*\widehat{D}_j(0,a)I(A_j=a),$$

$$\widehat{\mathsf{SED}}(z) = \frac{1}{J_1}\sum_{j=1}^{J}w_j^*\widehat{D}_j(z,1)I(A_j=1)-\frac{1}{J_0}\sum_{j=1}^{J}w_j^*\widehat{D}_j(z,0)I(A_j=0),$$

$$\widehat{\mathsf{DEY}}(a) = \frac{1}{J_a}\sum_{j=1}^{J}w_j^*\widehat{Y}_j(1,a)I(A_j=a)-\frac{1}{J_a}\sum_{j=1}^{J}w_j^*\widehat{Y}_j(0,a)I(A_j=a),$$

$$\widehat{\mathsf{SEY}}(z) = \frac{1}{J_1}\sum_{j=1}^{J}w_j^*\widehat{Y}_j(z,1)I(A_j=1)-\frac{1}{J_0}\sum_{j=1}^{J}w_j^*\widehat{Y}_j(z,0)I(A_j=0),$$

where

$$\widehat{D}_j(z,a) = \frac{\sum_{i=1}^{n_j}D_{ij}I(Z_{ij}=z)}{\sum_{i=1}^{n_j}I(Z_{ij}=z)}, \quad \widehat{Y}_j(z,a)=\frac{\sum_{i=1}^{n_j}Y_{ij}I(Z_{ij}=z)}{\sum_{i=1}^{n_j}I(Z_{ij}=z)}.$$

Theory of simple random sampling implies,

$$\mathbb{E}\{\widehat{D}_j(z,a) \mid A_j = a\} \;=\; \frac{1}{n_j}\sum_{i=1}^{n_j}\overline{D}_{ij}(z,a), \quad \mathbb{E}\{\widehat{Y}_j(z,a) \mid A_j = a\} \;=\; \frac{1}{n_j}\sum_{i=1}^{n_j}\overline{Y}_{ij}(z,a).$$

Thus, it is straightforward to show that the estimators for the ITT direct and spillover effects are unbiased. Below, without loss of generality, we prove the theorem only for the case with $w_j^* = 1$ since the general results can be obtained by simply transforming the outcome $Y_{ij}^* = w_j^* Y_{ij}$ and treatment receipt $D_{ij}^* = w_j^* D_{ij}$.

First, note that

$$\text{cov}(Z_{ij}, Z_{i'j}) = \begin{cases} \frac{n_{j1}}{n_j}\left(1 - \frac{n_{j1}}{n_j}\right) & \text{if } i = i', \\ -\frac{n_{j1}}{n_j(n_j-1)}\left(1 - \frac{n_{j1}}{n_j}\right) & \text{if } i \neq i'. \end{cases}$$

Then, from Assumption 1, we have

$$\begin{aligned}
& \text{var}\{\widehat{Y}_j(z,a) \mid A_j = a\} \\
= \;& \frac{1}{n_{jz}^2}\text{var}\left\{\sum_{i=1}^{n_j} Y_{ij} I(Z_{ij} = z) \mid A_j = a\right\} \\
= \;& \frac{1}{n_{jz}^2}\sum_{i=1}^{n_j} Y_{ij}^2(z,a)\text{cov}\{I(Z_{ij}=z), I(Z_{ij}=z)\} \\
& + \frac{1}{n_{jz}^2}\sum_{i \neq i'} Y_{ij}(z,a) Y_{i'j}(z,a)\text{cov}\{I(Z_{ij}=z), I(Z_{i'j}=z)\} \\
= \;& \frac{1}{n_j n_{jz}}\left(1 - \frac{n_{jz}}{n_j}\right)\sum_{i=1}^{n_j} Y_{ij}^2(z,a) - \frac{1}{n_{jz} n_j (n_j - 1)}\left(1 - \frac{n_{jz}}{n_j}\right)\sum_{i \neq i'} Y_{ij}(z,a) Y_{i'j}(z,a) \\
= \;& \frac{1}{n_j n_{jz}}\left(1 - \frac{n_{jz}}{n_j}\right)\sum_{i=1}^{n_j} Y_{ij}^2(z,a) - \frac{1}{n_{jz} n_j (n_j - 1)}\left(1 - \frac{n_{jz}}{n_j}\right)\left[\left\{\sum_{i=1}^{n_j} Y_{ij}(z,a)\right\}^2 - \sum_{i=1}^{n_j} Y_{ij}^2(z,a)\right] \\
= \;& \frac{1}{(n_j - 1) n_{jz}}\left(1 - \frac{n_{jz}}{n_j}\right)\sum_{i=1}^{n_j} Y_{ij}^2(z,a) - \frac{n_j}{n_{jz}(n_j - 1)}\left(1 - \frac{n_{jz}}{n_j}\right)\overline{Y}_j^2(z,a) \\
= \;& \frac{1}{n_{jz}}\left(1 - \frac{n_{jz}}{n_j}\right)\sigma_j^2(z,a),
\end{aligned}$$

and

$$\begin{aligned}
& \text{cov}\{\widehat{Y}_j(1,a), \widehat{Y}_j(0,a) \mid A_j = a\} \\
= \;& \frac{1}{n_{j1} n_{j0}}\sum_{i=1}^{n_j} Y_{ij}(1,a) Y_{ij}(0,a)\text{cov}\{I(Z_{ij}=1), I(Z_{ij}=0)\} \\
& + \frac{1}{n_{j1} n_{j0}}\sum_{i \neq i'} Y_{ij}(1,a) Y_{i'j}(0,a)\text{cov}\{I(Z_{ij}=1), I(Z_{i'j}=0)\} \\
= \;& -\frac{1}{n_j^2}\sum_{i=1}^{n_j} Y_{ij}(1,a) Y_{ij}(0,a) + \frac{1}{n_j^2(n_j - 1)}\sum_{i \neq i'} Y_{ij}(1,a) Y_{i'j}(0,a)
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{(n_j-1)n_j}\sum_{i=1}^{n_j} Y_{ij}(1,a)Y_{ij}(0,a) + \frac{1}{(n_j-1)}\overline{Y}_j(1,a)\overline{Y}_j(0,a) \\
&= -\frac{1}{n_j(n_j-1)}\sum_{i=1}^{n_j}\{Y_{ij}(1,a)-\overline{Y}_j(1,a)\}\{Y_{ij}(0,a)-\overline{Y}_j(0,a)\}.
\end{aligned}
$$

Therefore, we obtain,

$$
\begin{aligned}
&\text{var}\{\widehat{\mathsf{DEY}}_j(a) \mid A_j = a\} \\
&= \sum_{z=0,1} \text{var}\{\widehat{Y}_j(z,a) \mid A_j = a\} - 2\text{cov}\{\widehat{Y}_j(1,a), \widehat{Y}_j(0,a) \mid A_j = a\} \\
&= \sum_{z=0,1} \frac{1}{n_{jz}}\left(1-\frac{n_{jz}}{n_j}\right)\sigma_j^2(z,a) + \frac{1}{n_j}\left\{\sigma_j^2(1,a)+\sigma_j^2(0,a)-\omega_j^2(a)\right\} \\
&= \frac{\sigma_j^2(1,a)}{n_{j1}} + \frac{\sigma_j^2(0,a)}{n_{j0}} - \frac{\omega_j^2(a)}{n_j},
\end{aligned}
$$

which yields,

$$
\begin{aligned}
&\text{var}\{\widehat{\mathsf{DEY}}(a)\} \\
&= \mathbb{E}\left[\frac{1}{J_a^2}\sum_{j=1}^{J}\text{var}\{\widehat{\mathsf{DEY}}_j(a) \mid A_j = a\}I(A_j=a)\right] + \text{var}\left\{\frac{1}{J_a}\sum_{j=1}^{J}\mathsf{DEY}_j(a)I(A_j=a)\right\} \\
&= \left(1-\frac{J_a}{J}\right)\frac{\sigma_{DE}^2(a)}{J_a} + \frac{1}{J_a J}\sum_{j=1}^{J}\text{var}\left\{\widehat{\mathsf{DEY}}_j(a) \mid A_j = a\right\}.
\end{aligned}
$$

Next, we compute the variance of $\mathsf{SEY}(z)$. From Assumption 1, we have,

$$
\begin{aligned}
&\text{cov}\{\widehat{Y}(z,1), \widehat{Y}(z,0)\} \\
&= \frac{1}{J_1J_0}\text{cov}\left\{\sum_{j=1}^{J}\widehat{Y}_j(z,1)I(A_j=1), \sum_{j=1}^{J}\widehat{Y}_j(z,0)I(A_j=0)\right\} \\
&= \frac{1}{J_1J_0}\mathbb{E}\left[\text{cov}\left\{\sum_{j=1}^{J}\widehat{Y}_j(z,1)I(A_j=1), \sum_{j=1}^{J}\widehat{Y}_j(z,0)I(A_j=0) \mid A_1,\dots,A_J\right\}\right] \\
&\quad + \frac{1}{J_1J_0}\text{cov}\left\{\sum_{j=1}^{J}\overline{Y}_j(z,1)I(A_j=1), \sum_{j=1}^{J}\overline{Y}_j(z,0)I(A_j=0)\right\} \\
&= \frac{1}{J_1J_0}\text{cov}\left\{\sum_{j=1}^{J}\overline{Y}_j(z,1)I(A_j=1), \sum_{j=1}^{J}\overline{Y}_j(z,0)I(A_j=0)\right\} \\
&= \frac{1}{J_1J_0}\sum_{j=1}^{J}\overline{Y}_j(z,1)\overline{Y}_j(z,0)\text{cov}\{I(A_j=1), I(A_j=0)\} \\
&\quad + \frac{1}{J_1J_0}\sum_{j=j'}^{J}\overline{Y}_j(z,1)\overline{Y}_{j'}(z,0)\text{cov}\{I(A_j=1), I(A_{j'}=0)\} \\
&= -\frac{1}{J^2}\sum_{j=1}^{J}\overline{Y}_j(z,1)\overline{Y}_j(z,0) + \frac{1}{J^2(J-1)}\sum_{j=j'}^{J}\overline{Y}_j(z,1)\overline{Y}_{j'}(z,0)
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{J^2}\sum_{j=1}^{J}\overline{Y}_j(z,1)\overline{Y}_j(z,0) + \frac{1}{J-1}\overline{Y}(z,1)\overline{Y}(z,0) - \frac{1}{J^2(J-1)}\sum_{j=1}^{J}\overline{Y}_j(z,1)\overline{Y}_j(z,0) \\
&= -\frac{1}{J(J-1)}\sum_{j=1}^{J}\overline{Y}_j(z,1)\overline{Y}_j(z,0) + \frac{1}{J-1}\overline{Y}(z,1)\overline{Y}(z,0) \\
&= -\frac{1}{J(J-1)}\sum_{j=1}^{J}\{\overline{Y}_j(z,1)-\overline{Y}(z,1)\}\{\overline{Y}_j(z,0)-\overline{Y}(z,0)\} \\
&= \frac{1}{2J}\left\{\sigma^2_{SE}(z)-\sigma^2_b(z,1)-\sigma^2_b(z,0)\right\},
\end{aligned}
$$

where the second equality follows from the law of total variance and the third equality follows from the conditional independence $\boldsymbol{Z}_j \perp\!\!\!\perp \boldsymbol{Z}_{j'} \mid (A_1,\dots,A_J)$ for $j \neq j'$. Therefore, we have

$$
\begin{aligned}
& \text{var}\{\widehat{\mathsf{SEY}}(z)\} \\
=\ & \text{var}\{\widehat{Y}(z,1)\} + \text{var}\{\widehat{Y}(z,0)\} - 2\text{cov}\{\widehat{Y}(z,1),\widehat{Y}(z,0)\} \\
=\ & \left(1-\frac{J_1}{J}\right)\frac{\sigma^2_b(z,1)}{J_1} + \left(1-\frac{J_0}{J}\right)\frac{\sigma^2_b(z,0)}{J_0} - \frac{1}{J}\left\{\sigma^2_{SE}(z)-\sigma^2_b(z,1)-\sigma^2_b(z,0)\right\} \\
& + \frac{1}{J_1 J}\sum_{j=1}^{J}\text{var}\left\{\widehat{Y}_j(z,1)\mid A_j=1\right\} + \frac{1}{J_0 J}\sum_{j=1}^{J}\text{var}\left\{\widehat{Y}_j(z,0)\mid A_j=0\right\} \\
=\ & \frac{\sigma^2_b(z,1)}{J_1} + \frac{\sigma^2_b(z,0)}{J_0} - \frac{\sigma^2_{SE}(z)}{J} + \frac{1}{J_1 J}\sum_{j=1}^{J}\frac{1}{n_{jz}}\left(1-\frac{n_{jz}}{n_j}\right)\sigma^2_j(z,1) \\
& + \frac{1}{J_0 J}\sum_{j=1}^{J}\frac{1}{n_{jz}}\left(1-\frac{n_{jz}}{n_j}\right)\sigma^2_j(z,0).
\end{aligned}
$$

□

B.6 Covariances

B.6.1 cov($\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)$)

We introduce some notation. Define

$$
\begin{aligned}
\zeta_{jzz'}(a) &= \frac{1}{n_j-1}\sum_{i=1}^{n_j}\{Y_{ij}(z,a)-\overline{Y}_j(z,a)\}\{D_{ij}(z',a)-\overline{D}_j(z',a)\}, \\
\zeta_{j(1-0)}(a) &= \frac{1}{n_j-1}\sum_{i=1}^{n_j}\{\mathsf{DEY}_{ij}(a)-\mathsf{DEY}_j(a)\}\{\mathsf{DED}_{ij}(a)-\mathsf{DED}_j(a)\}.
\end{aligned}
$$

Similar to the calculation of var$\{\widehat{Y}_j(z,a) \mid A_j = a\}$ and cov$\{\widehat{Y}_j(1,a),\widehat{Y}_j(0,a) \mid A_j = a\}$ in the proof of Theorem 4, it is easy to show that

$$
\begin{aligned}
\text{cov}\{\widehat{Y}_j(1,a),\widehat{D}_j(0,a)\mid A_j=a\} &= -\frac{1}{n_j}\zeta_{j10}(a), \quad \text{cov}\{\widehat{Y}_j(0,a),\widehat{D}_j(1,a)\mid A_j=a\} = -\frac{1}{n_j}\zeta_{j01}(a), \\
\text{cov}\{\widehat{Y}_j(z,a),\widehat{D}_j(z,a)\mid A_j=a\} &= \frac{1}{n_{jz}}\left(1-\frac{n_{jz}}{n_j}\right)\zeta_{jzz}(a).
\end{aligned}
$$

Thus, we can obtain

$$
\begin{aligned}
\operatorname{cov}\left\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\right\} &= \sum_{z=0,1} \frac{1}{n_{jz}}\left(1-\frac{n_{jz}}{n_j}\right)\zeta_j(z,a) + \frac{1}{n_j}\zeta_{j10}(a) + \frac{1}{n_j}\zeta_{j01}(a) \\
&= \frac{\zeta_{j11}(a)}{n_{j1}} + \frac{\zeta_{j00}(a)}{n_{j0}} - \frac{\zeta_{j(1-0)}(a)}{n_j}.
\end{aligned}
$$

As a result,

$$
\begin{aligned}
\operatorname{cov}\left\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\right\} &= \mathbb{E}\left[\frac{1}{J_a^2}\sum_{j=1}^{J}\operatorname{cov}\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\} I(A_j = a)\right] \\
&\quad + \operatorname{cov}\left[\frac{1}{J_a}\sum_{j=1}^{J}\mathsf{DEY}_j(a) I(A_j = a), \frac{1}{J_a}\sum_{j=1}^{J}\mathsf{DED}_j(a) I(A_j = a)\right] \\
&= \left(1-\frac{J_a}{J}\right)\frac{\zeta_{DE}^2(a)}{J_a} + \frac{1}{J_a J}\sum_{j=1}^{J}\operatorname{cov}\left\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\right\}.
\end{aligned}
$$

B.6.2 cov($\widehat{\mathsf{SEY}}(a), \widehat{\mathsf{SED}}(a)$)

We introduce some notation. Define

$$
\begin{aligned}
\zeta_b^2(z,a) &= \frac{1}{J-1}\sum_{j=1}^{J}\{\overline{Y}_j(z,a) - \overline{Y}(z,a)\}\{\overline{D}_j(z,a) - \overline{D}(z,a)\}, \\
\zeta_{SE}^2(z) &= \frac{1}{J-1}\sum_{j=1}^{J}\{\mathsf{SEY}_j(z) - \mathsf{SEY}(z,a)\}\{\mathsf{SED}_j(z) - \mathsf{SEY}(z,a)\}.
\end{aligned}
$$

We can decompose $\operatorname{cov}\{\widehat{\mathsf{SEY}}(z), \widehat{\mathsf{SED}}(z)\}$ as,

$$
\begin{aligned}
& \operatorname{cov}\{\widehat{\mathsf{SEY}}(z), \widehat{\mathsf{SED}}(z)\} \\
= \; & \operatorname{cov}\{\widehat{Y}(z,1), \widehat{D}(z,1)\} + \operatorname{cov}\{\widehat{Y}(z,0), \widehat{D}(z,0)\} - \operatorname{cov}\{\widehat{Y}(z,1), \widehat{D}(z,0)\} - \operatorname{cov}\{\widehat{Y}(z,0), \widehat{Y}(z,0)\}.
\end{aligned}
$$

We then calculate each component,

$$
\begin{aligned}
\operatorname{cov}\{\widehat{Y}(z,a), \widehat{D}(z,a)\} &= \mathbb{E}\left[\frac{1}{J_a^2}\sum_{j=1}^{J}\operatorname{cov}\{\widehat{Y}_j(z,a), \widehat{D}_j(z,a) \mid A_j = a\} I(A_j = a)\right] \\
&\quad + \operatorname{cov}\left\{\frac{1}{J_a}\sum_{j=1}^{J}\overline{Y}_j(z,a) I(A_j = a), \frac{1}{J_a}\sum_{j=1}^{J}\overline{D}_j(z,a) I(A_j = a)\right\} \\
&= \left(1-\frac{J_a}{J}\right)\frac{\zeta_b^2(z,a)}{J_a} + \frac{1}{J_a J}\sum_{j=1}^{J}\operatorname{cov}\left\{\widehat{Y}_j(z,a), \widehat{D}_j(z,a) \mid A_j = a\right\},
\end{aligned}
$$

$$
\operatorname{cov}\{\widehat{Y}(z,1), \widehat{D}(z,0)\} = \operatorname{cov}\left\{\frac{1}{J_1}\sum_{j=1}^{J}\overline{Y}_j(z,1) I(A_j = 1), \frac{1}{J_0}\sum_{j=1}^{J}\overline{D}_j(z,0) I(A_j = 0)\right\}
$$

$$
\begin{aligned}
&= -\frac{1}{J(J-1)}\sum_{j=1}^{J}\{\overline{Y}_j(z,1)-\overline{Y}(z,1)\}\{\overline{D}_j(z,0)-\overline{D}(z,0)\}, \\
\text{cov}\{\widehat{Y}(z,0),\widehat{D}(z,1)\} &= \text{cov}\left\{\frac{1}{J_0}\sum_{j=1}^{J}\overline{Y}_j(z,0)I(A_j=0), \frac{1}{J_1}\sum_{j=1}^{J}\overline{D}_j(z,1)I(A_j=1)\right\} \\
&= -\frac{1}{J(J-1)}\sum_{j=1}^{J}\{\overline{Y}_j(z,0)-\overline{Y}(z,0)\}\{\overline{D}_j(z,1)-\overline{D}(z,1)\}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \text{cov}\{\widehat{Y}(z,1),\widehat{D}(z,0)\}+\text{cov}\{\widehat{Y}(z,0),\widehat{D}(z,1)\} \\
= \; & -\frac{1}{J(J-1)}\sum_{j=1}^{J}\{\overline{Y}_j(z,0)-\overline{Y}(z,0)\}\{\overline{D}_j(z,1)-\overline{D}(z,1)\} \\
& -\frac{1}{J(J-1)}\sum_{j=1}^{J}\{\overline{Y}_j(z,1)-\overline{Y}(z,1)\}\{\overline{D}_j(z,0)-\overline{D}(z,0)\} \\
= \; & \frac{1}{J}\{\zeta_{SE}^2-\zeta_b^2(z,1)-\zeta_b^2(z,0)\}.
\end{aligned}
$$

As a result, we have

$$
\begin{aligned}
& \text{cov}\{\widehat{\mathsf{SEY}}(z),\widehat{\mathsf{SED}}(z)\} \\
= \; & \left(1-\frac{J_1}{J}\right)\frac{\zeta_b^2(z,1)}{J_1}+\frac{1}{J_1J}\sum_{j=1}^{J}\text{cov}\left\{\widehat{Y}_j(z,1),\widehat{D}_j(z,1)\mid A_j=1\right\}+\left(1-\frac{J_0}{J}\right)\frac{\zeta_b^2(z,0)}{J_a} \\
& +\frac{1}{J_0J}\sum_{j=1}^{J}\text{cov}\left\{\widehat{Y}_j(z,0),\widehat{D}_j(z,0)\mid A_j=0\right\}-\frac{1}{J}\{\zeta_{SE}^2-\zeta_b^2(z,1)-\zeta_b^2(z,0)\} \\
= \; & \frac{\zeta_b^2(z,1)}{J_1}+\frac{\zeta_b^2(z,0)}{J_0}-\frac{\zeta_{SE}^2(z)}{J}+\frac{1}{J_1J}\sum_{j=1}^{J}\text{cov}\left\{\widehat{Y}_j(z,1),\widehat{D}_j(z,1)\mid A_j=1\right\} \\
& +\frac{1}{J_0J}\sum_{j=1}^{J}\text{cov}\left\{\widehat{Y}_j(z,0),\widehat{D}_j(z,0)\mid A_j=0\right\}.
\end{aligned}
$$

B.7 Variance Estimators for the ITT Effects

We first show that the variance estimator is conservative for var$\{\widehat{\mathsf{DEY}}(a)\}$. From the classical theory of simple random sampling, we know $\mathbb{E}\{\widehat{\sigma}_j^2(z,a)\mid A_j=a\}=\sigma_j^2(z,a)$. In addition, we have

$$
\begin{aligned}
& \mathbb{E}\{\widehat{\sigma}_{DE}^2(a)\} \\
= \; & \frac{1}{J_a-1}\mathbb{E}\left\{\sum_{j=1}^{J}\widehat{\mathsf{DEY}}_j^2(a)I(A_j=a)-J_a\widehat{\mathsf{DEY}}(a)^2\right\} \\
= \; & \frac{1}{J_a-1}\mathbb{E}\left(\sum_{j=1}^{J}[\text{var}\{\widehat{\mathsf{DEY}}_j(a)\mid A_j=a\}+\mathsf{DEY}_j(a)^2]I(A_j=a)\right)
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{J_a}{J_a-1}[\text{var}\{\widehat{\mathsf{DEY}}(a)\} + \mathsf{DEY}(a)^2] \\
= \;& \frac{J_a}{J(J_a-1)}\sum_{j=1}^{J}[\text{var}\{\widehat{\mathsf{DEY}}_j(a) \mid A_j = a\} + \frac{J_a(J-1)}{(J_a-1)J}\sigma^2_{DE}(a) - \frac{J_a}{J_a-1}\text{var}\{\widehat{\mathsf{DEY}}(a)\} \\
= \;& \sigma^2_{DE}(a) + \frac{1}{J}\sum_{j=1}^{J}[\text{var}\{\widehat{\mathsf{DEY}}_j(a) \mid A_j = a\}.
\end{aligned}
$$

and

$$
\begin{aligned}
& \mathbb{E}\{\widehat{\sigma}^2_b(z,a)\} \\
= \;& \frac{1}{J_a-1}\mathbb{E}\left\{\sum_{j=1}^{J}\widehat{Y}^2_j(z,a)I(A_j=a) - J_a\widehat{Y}(z,a)^2\right\} \\
= \;& \frac{1}{J_a-1}\mathbb{E}\left(\sum_{j=1}^{J}[\text{var}\{\widehat{Y}_j(z,a) \mid A_j=a\} + \overline{Y}_j(z,a)^2]I(A_j=a)\right) - \frac{J_a}{J_a-1}[\text{var}\{\widehat{Y}(z,a)\} + \overline{Y}(z,a)^2] \\
= \;& \frac{J_a}{J(J_a-1)}\sum_{j=1}^{J}[\text{var}\{\widehat{Y}_j(z,a) \mid A_j=a\} + \frac{J_a}{J(J_a-1)}\sum_{j=1}^{J}\overline{Y}_j(z,a)^2 \\
& -\frac{J_a}{J_a-1}[\text{var}\{\widehat{Y}(z,a)\} + \overline{Y}(z,a)^2] \\
= \;& \frac{J_a}{J(J_a-1)}\sum_{j=1}^{J}[\text{var}\{\widehat{Y}_j(z,a) \mid A_j=a\} + \frac{J_a(J-1)}{(J_a-1)J}\sigma^2_b(z,a) - \frac{J_a}{J_a-1}\text{var}\{\widehat{Y}(z,a)\} \\
= \;& \frac{J_a}{J(J_a-1)}\sum_{j=1}^{J}[\text{var}\{\widehat{Y}_j(z,a) \mid A_j=a\} + \frac{J_a(J-1)}{(J_a-1)J}\sigma^2_b(z,a) \\
& -\frac{J_a}{J_a-1}\left[\left(1-\frac{J_a}{J}\right)\frac{\sigma^2_b(z,a)}{J_a} + \frac{1}{J_aJ}\sum_{j=1}^{J}\text{var}\left\{\widehat{Y}_j(z,1) \mid A_j=a\right\}\right] \\
= \;& \sigma^2_b(z,a) + \frac{1}{J}\sum_{j=1}^{J}[\text{var}\{\widehat{Y}_j(z,a) \mid A_j=a\}.
\end{aligned}
$$

Therefore, we have,

$$
\begin{aligned}
& \mathbb{E}[\widehat{\text{var}}\{\widehat{\mathsf{DEY}}(a)\}] \\
= \;& \left(1-\frac{J_a}{J}\right)\frac{\sigma^2_{DE}(a)}{J_a} + \left(1-\frac{J_a}{J}\right)\frac{1}{J_aJ}\sum_{j=1}^{J}\text{var}\left\{\widehat{\mathsf{DEY}}_j(a) \mid A_j=a\right\} \\
& +\frac{1}{J^2}\sum_{j=1}^{J}\left\{\frac{\sigma^2_j(1,a)}{n_{j1}} + \frac{\sigma^2_j(0,a)}{n_{j0}}\right\} \\
= \;& \text{var}\{\widehat{\mathsf{DEY}}(a)\} + \frac{1}{J^2}\sum_{j=1}^{J}\frac{\omega^2_j(a)}{n_{j1}} \\
\geq \;& \text{var}\{\widehat{\mathsf{DEY}}(a)\}.
\end{aligned}
$$

We then consider the variance estimator for $\mathsf{SEY}(z)$,

$$\widehat{\text{var}}\left\{\widehat{\mathsf{SEY}}(z)\right\} = \frac{\widehat{\sigma}_b^2(z,1)}{J_1} + \frac{\widehat{\sigma}_b^2(z,0)}{J_0}.$$

We have

$$\begin{aligned}
&\mathbb{E}[\widehat{\text{var}}\{\widehat{\mathsf{SEY}}(z)\}] \\
= \; & \frac{\sigma_b^2(z,1)}{J_1} + \frac{\sigma_b^2(z,0)}{J_0} + \sum_{a=0}^{1} \frac{1}{J_a J} \sum_{j=1}^{J} \text{var}\left\{\widehat{Y}_j(z,a) \mid A_j = a\right\} \\
= \; & \text{var}\{\widehat{\mathsf{SEY}}(z)\} + \frac{\sigma_{SE}^2(z)}{J} \\
\geq \; & \text{var}\{\widehat{\mathsf{SEY}}(z)\}.
\end{aligned}$$

Next, we consider the estimator for $\text{cov}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\}$. Similarly, we have $\mathbb{E}\{\widehat{\zeta}_j(z,a) \mid A_j = a\} = \zeta_j(z,a)$ and hence,

$$\begin{aligned}
&\mathbb{E}\{\widehat{\zeta}_{DE}^2(a)\} \\
= \; & \frac{1}{J_a - 1}\mathbb{E}\left\{\sum_{j=1}^{J} \widehat{\mathsf{DEY}}_j(a)\widehat{\mathsf{DED}}_j(a) I(A_j = a) - J_a \widehat{\mathsf{DEY}}(a)\widehat{\mathsf{DED}}(a)\right\} \\
= \; & \frac{1}{J_a - 1}\mathbb{E}\left(\sum_{j=1}^{J} \left[\text{cov}\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\} + \mathsf{DEY}_j(a)\mathsf{DED}_j(a)\right] I(A_j = a)\right) \\
& - \frac{J_a}{J_a - 1}\left[\text{cov}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\} + \mathsf{DEY}(a)\mathsf{DED}(a)\right] \\
= \; & \frac{J_a}{J(J_a - 1)} \sum_{j=1}^{J} [\text{cov}\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\} + \frac{J_a(J-1)}{(J_a - 1)J}\zeta_{DE}(a) \\
& - \frac{J_a}{J_a - 1}\text{cov}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\} \\
= \; & \zeta_{DE}^2(a) + \frac{1}{J}\sum_{j=1}^{J} [\text{cov}\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\}.
\end{aligned}$$

Therefore, we have,

$$\begin{aligned}
&\mathbb{E}[\widehat{\text{cov}}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\}] \\
= \; & \left(1 - \frac{J_a}{J}\right)\frac{\zeta_{DE}^2(a)}{J_a} + \left(1 - \frac{J_a}{J}\right)\frac{1}{J_a J}\sum_{j=1}^{J} \text{cov}\left\{\widehat{\mathsf{DEY}}_j(a), \widehat{\mathsf{DED}}_j(a) \mid A_j = a\right\} \\
& + \frac{1}{J^2}\sum_{j=1}^{J}\left\{\frac{\zeta_j(1,a)}{n_{j1}} + \frac{\zeta_j(0,a)}{n_{j0}}\right\} \\
= \; & \text{cov}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\} + \frac{1}{J^2}\sum_{j=1}^{J} \frac{\zeta_{j(1-0)}(a)}{n_{j1}} \\
\geq \; & \text{cov}\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\}.
\end{aligned}$$

B.8 Asymptotically Conservative Variance Estimator for the CADE

Although $\widehat{\text{var}}\left\{\widehat{\mathsf{CADE}}(a)\right\}$ in equation (9) is not a conservative variance estimator in finite samples, we show that it is asymptotically conservative. First, the asymptotic variance of $\widehat{\mathsf{CADE}}(a)$ can be rewritten as,

$$\text{var}\left[\frac{1}{\mathsf{DED}(a)}\left\{\widehat{\mathsf{DEY}}(a) - \mathsf{CADE}(a)\cdot\widehat{\mathsf{DED}}(a)\right\}\right],$$

which is the variance of a linear combination of $\widehat{\mathsf{DED}}(a)$ and $\widehat{\mathsf{DEY}(a)}$. Similar to the proof in Section B.7, we can show,

$$\begin{aligned}&\mathbb{E}\left\{\frac{1}{\mathsf{DED}(a)^2}\left[\widehat{\text{var}}\left\{\widehat{\mathsf{DEY}}(a)\right\} - 2\frac{\mathsf{DEY}(a)}{\mathsf{DED}(a)}\widehat{\text{cov}}\left\{\widehat{\mathsf{DEY}}(a), \widehat{\mathsf{DED}}(a)\right\} + \frac{\mathsf{DEY}(a)^2}{\mathsf{DED}(a)^2}\widehat{\text{var}}\left\{\widehat{\mathsf{DED}}(a)\right\}\right]\right\}\\ \geq\ & \text{var}\left[\frac{1}{\mathsf{DED}(a)}\left\{\widehat{\mathsf{DEY}}(a) - \mathsf{CADE}(a)\cdot\widehat{\mathsf{DED}}(a)\right\}\right].\end{aligned}$$

Under the restriction on interference in (Sävje *et al.*, 2017), $\widehat{\mathsf{DED}}(a)$ converges to $\mathsf{DED}(a)$ and $\widehat{\mathsf{DEY}}(a)$ converges to $\mathsf{DEY}(a)$. Therefore, we obtain the desired result for a bounded outcome,

$$\mathbb{E}\left[\widehat{\text{var}}\left\{\widehat{\mathsf{CADE}}(a)\right\}\right] \geq \text{avar}\left\{\widehat{\mathsf{CADE}}(a)\right\}.$$

□

C Proofs for the Regression-Based Approach

As in case of the randomization inference approach, it is suffice to prove the case with $w_j^* = 1$ since the results are applicable directly with any other weight by multiplying D_{ij} and Y_{ij} with appropriate constants. Because the columns in the design matrix of the regression models corresponding to different treatment assignment mechanisms are orthogonal to each other, we can prove the results separately for each treatment assignment mechanism. Therefore, we prove the theorems for a given a and with abuse of notation use the same notation for the sub-matrix that consists of the columns corresponding to the treatment assignment mechanism a in a full matrix. For example, the proof of Theorem 6 uses $\mathbf{X}_{ij}$ to denote $(I(A_j = a), Z_{ij}I(A_j = a))$ while in the main text we use $\mathbf{X}_{ij}$ to represent $(I(A_j = 1), I(A_j = 0), Z_{ij}I(A_j = 1), Z_{ij}I(A_j = 0))$.

C.1 Proof of Theorem 6

Define,

$$N_{za} \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j} I(Z_{ij} = z, A_j = a)w_{ij} \;=\; \sum_{j=1}^{J}\sum_{i=1}^{n_j} I(Z_{ij} = z, A_j = a)\frac{1}{J_a n_{jz}} \;=\; 1$$

and $N_{+a} = N_{0a} + N_{1a}$. Then, the OLS estimate can be written as,

$$(\widehat{\alpha}_a, \widehat{\alpha}_{1a})^\top \;=\; (\mathbf{X}^\top\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{W}\boldsymbol{Y},$$

where

$$
\begin{aligned}
(\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} &= \begin{pmatrix} N_{+a} & N_{1a} \\ N_{1a} & N_{1a} \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \\
\mathbf{X}^\top \mathbf{W} \boldsymbol{Y} &= \begin{pmatrix} \widehat{Y}(1,a) + \widehat{Y}(0,a) \\ \widehat{Y}(1,a) \end{pmatrix}.
\end{aligned}
$$

Therefore, we have,

$$
\widehat{\alpha}_a = \widehat{Y}(0,a), \quad \widehat{\alpha}_{1a} = \widehat{\mathsf{DEY}}(a).
$$

□

C.2 Proof of Theorem 7

We first write the two-stage least squares regression as,

$$
\begin{aligned}
Y_{ij} &= \sum_{a=0}^{1} \beta_a I(A_j = a) + \sum_{a=0}^{1} \beta_{1a} \widehat{D}_{ij} I(A_j = a) + \epsilon_{ij}, \\
\widehat{D}_{ij} &= \sum_{a=0}^{1} \widehat{\gamma}_a^{\text{wls}} I(A_j = a) + \sum_{a=0}^{1} \widehat{\gamma}_{1a}^{\text{wls}} Z_{ij} I(A_j = a).
\end{aligned}
$$

where $\widehat{\gamma}_a^{\text{wls}}$ and $\widehat{\gamma}_{1a}^{\text{wls}}$ are the weighted least squares estimate of the corresponding coefficients from the model given in equation (10). Then, we obtain,

$$
\begin{aligned}
Y_{ij} &= \sum_{a=0}^{1} \beta_a I(A_j = a) + \sum_{a=0}^{1} \beta_{1a} \left\{ \sum_{a=0}^{1} \widehat{\gamma}_a^{\text{wls}} I(A_j = a) + \sum_{a=0}^{1} \widehat{\gamma}_{1a}^{\text{wls}} Z_{ij} I(A_j = a) \right\} I(A_j = a) + \epsilon_{ij} \\
&= \sum_{a=0}^{1} \beta_a I(A_j = a) + \sum_{a=0}^{1} \beta_{1a} \left\{ \widehat{\gamma}_a^{\text{wls}} I(A_j = a) + \widehat{\gamma}_{1a}^{\text{wls}} Z_{ij} I(A_j = a) \right\} + \epsilon_{ij} \\
&= \sum_{a=0}^{1} (\beta_a + \beta_{1a} \widehat{\gamma}_a^{\text{wls}}) I(A_j = a) + \sum_{a=0}^{1} \beta_{1a} \widehat{\gamma}_{1a}^{\text{wls}} Z_{ij} I(A_j = a) + \epsilon_{ij}.
\end{aligned}
$$

Comparison of this with the weighted regression model of Y_{ij} on Z_{ij} given in equation (11) implies,

$$
\widehat{\alpha}_a^{\text{wls}} = \widehat{\beta}_a^{\text{w2sls}} + \widehat{\beta}_{1a}^{\text{w2sls}} \widehat{\gamma}_a^{\text{wls}}, \quad \widehat{\alpha}_{1a}^{\text{wls}} = \widehat{\beta}_{1a}^{\text{w2sls}} \widehat{\gamma}_{1a}^{\text{wls}}.
$$

Thus, using Theorem 6, we have,

$$
\widehat{\beta}_{1a}^{\text{w2sls}} = \widehat{\mathsf{CADE}}(a), \quad \widehat{\beta}_a^{\text{w2sls}} = \widehat{Y}(0,a) - \widehat{\mathsf{CADE}}(a) \cdot \widehat{D}(0,a).
$$

C.3 Proof of Theorem 8

We prove the results only for the direct effects on the outcome. The results for the direct effects on the treatment receipt are similar.

$$
\mathbf{P}_j = \mathbf{W}_j^{1/2} \mathbf{X}_j (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}_j^\top \mathbf{W}_j^{1/2}
$$

$$
\begin{aligned}
&= \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}}\mathbf{1}_{n_{j0}} & \mathbf{0}_{n_{j0}} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}}\mathbf{1}_{n_{j0}} & \mathbf{0}_{n_{j0}} \end{pmatrix}^\top \\
&= \begin{pmatrix} \mathbf{0}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}}\mathbf{1}_{n_{j0}} & -\frac{1}{\sqrt{J_a n_{j0}}}\mathbf{1}_{n_{j0}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}}\mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}}\mathbf{1}_{n_{j0}} & \mathbf{0}_{n_{j0}} \end{pmatrix}^\top \\
&= \begin{pmatrix} \frac{1}{J_a n_{j1}}\mathbf{1}_{n_{j1}\times n_{j1}} & \mathbf{0}_{n_{j1}\times n_{j0}} \\ \mathbf{0}_{n_{j0}\times n_{j1}} & \frac{1}{J_a n_{j0}}\mathbf{1}_{n_{j0}\times n_{j0}} \end{pmatrix},
\end{aligned}
$$

where $\mathbf{1}_m$ ($\mathbf{0}_m$) is an m-dimensional vector of ones (zeros) and $\mathbf{1}_{m_1 m_2}$ ($\mathbf{0}_{m_1 m_2}$) is an $m_1 \times m_2$ dimensional matrix of ones (zeros).

Since $(\mathbf{1}_{n_{j1}}^\top, \mathbf{0}_{n_{j0}}^\top)^\top$ and $(\mathbf{0}_{n_{j1}}^\top, \mathbf{1}_{n_{j0}}^\top)^\top$ are two eigenvectors of $\boldsymbol{I}_{n_j} - \mathbf{P}_j$ whose eigenvalue is $(J_a - 1)/J_a$, we have,

$$
\begin{aligned}
(\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2}(\mathbf{1}_{n_{j1}}^\top, \mathbf{0}_{n_{j0}}^\top)^\top &= \sqrt{\frac{J_a}{J_a - 1}}(\mathbf{1}_{n_{j1}}^\top, \mathbf{0}_{n_{j0}}^\top)^\top, \\
(\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2}(\mathbf{0}_{n_{j1}}^\top, \mathbf{1}_{n_{j0}}^\top)^\top &= \sqrt{\frac{J_a}{J_a - 1}}(\mathbf{0}_{n_{j1}}^\top, \mathbf{1}_{n_{j0}}^\top)^\top.
\end{aligned}
$$

Thus,

$$
(\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2}\mathbf{W}_j\mathbf{X}_j = \sqrt{\frac{J_a}{J_a - 1}} \begin{pmatrix} \frac{1}{J_a n_{j1}}\mathbf{1}_{n_{j1}} & \frac{1}{J_a n_{j1}}\mathbf{1}_{n_{j1}} \\ \frac{1}{J_a n_{j0}}\mathbf{1}_{n_{j0}} & \mathbf{0}_{n_{j0}} \end{pmatrix}.
$$

For a unit with $(A_j = a, Z_{ij} = 1)$, we have $\widehat{\epsilon}_{ij} = Y_{ij} - \widehat{\alpha}_a - \widehat{\alpha}_{1a} = Y_{ij} - \widehat{Y}(1, a)$, and for a unit with $(A_j = a, Z_{ij} = 0)$, we have $\widehat{\epsilon}_{ij} = Y_{ij} - \widehat{\alpha}_a = Y_{ij} - \widehat{Y}(0, a)$. As a result,

$$
\begin{aligned}
&\widehat{\boldsymbol{\epsilon}}_j^\top (\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2}\mathbf{W}_j\mathbf{X}_j \\
&= \sqrt{\frac{J_a}{J_a - 1}}(Y_{1j} - \widehat{Y}(1, a), \ldots, Y_{n_j j} - \widehat{Y}(0, a)) \begin{pmatrix} \frac{1}{J_a n_{j1}}\mathbf{1}_{n_{j1}} & \frac{1}{J_a n_{j1}}\mathbf{1}_{n_{j1}} \\ \frac{1}{J_a n_{j0}}\mathbf{1}_{n_{j0}} & \mathbf{0}_{n_{j0}} \end{pmatrix}, \\
&= \sqrt{\frac{J_a}{J_a - 1}} \begin{pmatrix} \frac{1}{J_a n_{j1}} \left\{ \sum_{i=1}^{n_j} Y_{ij} Z_{ij} - n_{j1}\widehat{Y}(1, 1) \right\} + \frac{1}{J_a n_{j0}} \left\{ \sum_{i=1}^{n_j} Y_{ij}(1 - Z_{ij}) - n_{j0}\widehat{Y}(0, 1) \right\} \\ \frac{1}{J_a n_{j1}} \left\{ \sum_{i=1}^{n_j} Y_{ij} Z_{ij} - n_{j1}\widehat{Y}(1, 1) \right\} \end{pmatrix}^\top \\
&= \sqrt{\frac{1}{J_a(J_a - 1)}} \begin{pmatrix} \widehat{Y}_j(1, 1) - \widehat{Y}(1, 1) + \widehat{Y}_j(0, 1) - \widehat{Y}(0, 1) \\ \widehat{Y}_j(1, 1) - \widehat{Y}(1, 1) \end{pmatrix}^\top.
\end{aligned}
$$

Let $\mathbf{V}_j = \mathbf{X}_j^\top \mathbf{W}_j (\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2} \widehat{\boldsymbol{\epsilon}}_j \widehat{\boldsymbol{\epsilon}}_j^\top (\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2}\mathbf{W}_j\mathbf{X}_j$ and $v_{jk_1k_2}$ be the k_1k_2-th entry of $\mathbf{V}_j$. Then,

$$
\begin{aligned}
v_{j11} &= \frac{1}{J_a(J_a - 1)} \left\{ \widehat{Y}_j(1, a) - \widehat{Y}(1, a) + \widehat{Y}_j(0, a) - \widehat{Y}(0, a) \right\}^2, \\
v_{j12} &= v_{j22} + \frac{1}{J_a(J_a - 1)} \left\{ \widehat{Y}_j(1, a) - \widehat{Y}(1, a) \right\} \left\{ \widehat{Y}_j(0, a) - \widehat{Y}(0, a) \right\}, \\
v_{j22} &= \frac{1}{J_a(J_a - 1)} \left\{ \widehat{Y}_j(1, a) - \widehat{Y}(1, a) \right\}^2.
\end{aligned}
$$

The definition of the cluster-robust HC2 variance implies,

$$\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}\{(\widehat{\alpha}_a^{\text{wls}}, \widehat{\alpha}_{1a}^{\text{wls}})^\top\} \;=\; (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \left\{ \sum_{j=1}^{J} \mathbf{V}_j I(A_j = a) \right\} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1},$$

which yields

$$\begin{aligned}
\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_{1a}^{\text{wls}}) &= \sum_{j=1}^{J} (4v_{j22} - 2v_{j12} - 2v_{j21} + v_{j11}) I(A_j = a) \\
&= \frac{1}{J_a(J_a - 1)} \sum_{j=1}^{J} [\{\widehat{Y}_j(1,a) - \widehat{Y}_j(0,a)\} - \{\widehat{Y}(1,a) - \widehat{Y}(0,a)\}]^2 \\
&= \frac{\widehat{\sigma}_{DE}^2(a)}{J_a}, \\
\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_{1}^{\text{wls}}) &= \sum_{j=1}^{J} (v_{j22} - v_{j12} - v_{j21} + v_{j11}) I(A_j = a) \\
&= \frac{1}{J_a(J_a - 1)} \sum_{j=1}^{J} \{\widehat{Y}_j(0,a) - \widehat{Y}(0,a)\}^2 \\
&= \frac{\widehat{\sigma}_b^2(0,a)}{J_a},
\end{aligned}$$

and

$$\begin{aligned}
\widehat{\text{cov}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_a^{\text{wls}}, \widehat{\alpha}_{1a}^{\text{wls}}) &= \frac{1}{J_a^2} \sum_{j} (-2v_{j22} + 2v_{j12} + v_{j21} - v_{j11}) I(A_j = a) \\
&= \frac{1}{J_a(J_a - 1)} \left\{ \widehat{Y}_j(1,a) - \widehat{Y}(1,a) \right\} \left\{ \widehat{Y}_j(0,a) - \widehat{Y}(0,a) \right\} \\
&\quad - \frac{1}{J_a(J_a - 1)} \sum_{j=1}^{J} \{\widehat{Y}_j(0,a)\} - \widehat{Y}(0,a)\}^2.
\end{aligned}$$

Thus, we obtain,

$$\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_a^{\text{wls}} + \widehat{\alpha}_{1a}^{\text{wls}}) \;=\; \widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_{1a}^{\text{wls}}) + \widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_{11}^{\text{wls}}) + 2\widehat{\text{cov}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_a^{\text{wls}}, \widehat{\alpha}_{1a}^{\text{wls}}) \;=\; \frac{\widehat{\sigma}_b^2(1,a)}{J_a}.$$

Next, we calculate the individual-robust HC2 variance. Similarly, using the orthogonality among the covariates, we have

$$\widehat{\text{var}}_{\text{hc2}}^{\text{ind}}\{(\widehat{\alpha}_a^{\text{wls}}, \widehat{\alpha}_{1a}^{\text{wls}})^\top\} \;=\; (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \left\{ \sum_{j=1}^{J} \sum_{i=1}^{n_j} w_{ij}^2 \widehat{\epsilon}_{ij}^{*2} I(A_j = a)(1 - P_{ij})^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right\} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1},$$

For units with $Z_{ij} = z$, we obtain

$$\begin{aligned}
P_{ij} &= w_{ij} J_a \mathbf{X}_{ij}^\top \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \mathbf{X}_{ij} \;=\; \frac{1}{n_{jz}}, \\
\widehat{\epsilon}_{ij}^{*} &= Y_{ij} - \widehat{Y}_j(z,a).
\end{aligned}$$

Therefore, we have

$$
\begin{aligned}
& \sum_{j=1}^{J}\sum_{i=1}^{n_j} w_{ij}^2 \widehat{\epsilon}_{ij}^{*2} I(A_j=a)(1-P_{ij})^{-1}\mathbf{X}_{ij}\mathbf{X}_{ij}^\top \\
= & \frac{1}{J_a^2}\sum_{j=1}^{J} I(A_j=a)\sum_{i=1}^{n_j}\left\{\frac{(Y_{ij}-\widehat{Y}_j(1,a))^2 I(Z_{ij}=1)}{n_{j1}(n_{j1}-1)} + \frac{(Y_{ij}-\widehat{Y}_j(0,a))^2 I(Z_{ij}=0)}{n_{j0}(n_{j0}-1)}\right\}\mathbf{X}_{ij}\mathbf{X}_{ij}^\top \\
= & \frac{1}{J_a^2}\begin{pmatrix} \sum_{j=1}^{J}\left(\frac{\widehat{\sigma}_j^2(1,a)}{n_{j1}} + \frac{\widehat{\sigma}_j^2(0,a)}{n_j-n_{j1}}\right) I(A_j=a) & \sum_{j=1}^{J}\frac{\widehat{\sigma}_j^2(1,a)}{n_{j1}} I(A_j=a) \\ \sum_{j=1}^{J}\frac{\widehat{\sigma}_j^2(1,a)}{n_{j1}} I(A_j=a) & \sum_{j=1}^{J}\frac{\widehat{\sigma}_j^2(0,a)}{n_j-n_{j1}} I(A_j=a)\end{pmatrix}.
\end{aligned}
$$

As a result,

$$
\widehat{\text{var}}_{\text{hc2}}^{\text{ind}}(\widehat{\alpha}_{11}^{\text{wls}}) = \frac{1}{J_a^2}\sum_{j=1}^{J}\left(\frac{\widehat{\sigma}_j^2(1,a)}{n_{j1}} + \frac{\widehat{\sigma}_j^2(0,a)}{n_j-n_{j1}}\right) I(A_j=a),
$$

and hence,

$$
\widehat{\text{var}}\left\{\widehat{\mathsf{DEY}}(a)\right\} = \left(1-\frac{J_a}{J}\right)\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\alpha}_{1a}^{\text{wls}}) + \frac{J_a}{J}\widehat{\text{var}}_{\text{hc2}}^{\text{ind}}(\widehat{\alpha}_{1a}^{\text{wls}}).
$$

□

C.4 Relations to Random Effects Models for Spilt-Plot Designs

Because two-stage experiments have a hierarchical structure, we re-express the linear model as a random effects model. First, suppose that $Y_{ij}(1,a) - Y_{ij}(0,a) = \alpha_{aij}$, then we can write the potential outcomes as,

$$
Y_{ij}(z,a) = I(z=1)\overline{\alpha}_{aj} + \overline{\alpha}_{0aj} + I(z=1)r_{aij} + r_{0aij},
$$

where $\overline{\alpha}_{0aj} = \overline{Y}_j(0,a)$, $\overline{\alpha}_{aj} = Y_j(1,a) - Y_j(0,a)$, $r_{0aij} = Y_{ij}(0,a) - \overline{Y}_j(0,a)$ and $r_{aij} = \alpha_{aij} - \overline{\alpha}_{aj}$. Then, the realized outcome can be expressed as,

$$
\begin{aligned}
Y_{ij} = & \sum_{a=0,1}\overline{\alpha}_{0aj} I(A_j=a) + \sum_{a=0,1}\alpha_{aj} Z_{ij} I(A_j=a) + \sum_{a=0,1} I(A_j=a) r_{0aij} + \sum_{a=0,1} Z_{ij} I(A_j=a) r_{aij}, \\
= & \sum_{a=0,1}\overline{\alpha}_{0a} I(A_j=a) + \sum_{a=0,1}\overline{\alpha}_a Z_{ij} I(A_j=a) + \sum_{a=0,1} I(A_j=a) r_{0aij} + \sum_{a=0,1} Z_{ij} I(A_j=a) r_{aij} \\
& + \sum_{a=0,1} s_{0aj} I(A_j=a) + \sum_{a=0,1} s_{aj} Z_{ij} I(A_j=a),
\end{aligned} \tag{A13}
$$

where $\overline{\alpha}_a = \overline{Y}(1,a) - \overline{Y}(0,a)$, $\overline{\alpha}_{0a} = \overline{Y}(0,a)$, $s_{aj} = \overline{\alpha}_{aj} - \overline{\alpha}_a$ and $s_{0aj} = \overline{\alpha}_{0aj} - \overline{\alpha}_{0a}$.

Given the similarity between equations (A13) and (11), we can treat the last four terms of equation (A13) as the error term ϵ_{ij} in equation (11), and decompose it into two parts, i.e., $\epsilon_{ij} = \epsilon_{Bj} + \epsilon_{Wij}$, where

$$
\begin{aligned}
\epsilon_{Bj} &= \sum_{a=0,1} s_{0aj} I(A_j=a) + \sum_{a=0,1} s_{aj} Z_{ij} I(A_j=a) = \sum_{a=0,1} I(A_j=a)(s_{0aj} + s_{aj} Z_{ij}), \\
\epsilon_{Wij} &= \sum_{a=0,1} I(A_j=a) r_{0aij} + \sum_{a=0,1} Z_{ij} I(A_j=a) r_{aij} = \sum_{a=0,1} I(A_j=a)(r_{0aij} + r_{aij} Z_{ij}),
\end{aligned}
$$

where ϵ_{Bj} can be viewed as the between-cluster residual and ϵ_{Wij} can be viewed as the within-cluster residual. The cluster-robust HC2 variance in our regression-based approach corresponds to ϵ_{Bj} and the individual-robust HC2 variance corresponds to ϵ_{Wij}. Because $\sum_{i=1}^{n_j} \epsilon_{Wij} = \sum_{i=1}^{n_j} \epsilon_{Wij} I(Z_{ij} = z) = 0$ holds for all j and $z = 0, 1$, the adjustment for $\widehat{\epsilon}^*_{ij}$ is necessary to ensure the residual from our regression also satisfies this condition. The term $w_{ij}^2 \widehat{\epsilon}_{ij}^{*2} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top$ in the individual-robust variance corresponds to $r_{aij} + Z_{ij} r_{0aij}$, and the term $\mathbf{X}_j^\top \boldsymbol{W}_j \widehat{\boldsymbol{\epsilon}}_j \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{W}_j \boldsymbol{X}_j$ in the cluster-robust variance corresponds to $s_{0aj} + s_{aj} Z_{ij}$.

C.5 Proof of Theorem 9

Define

$$
\begin{aligned}
\widehat{\xi}_j^2(z, a) &= \frac{\sum_{i=1}^{n_j} \{D_{ij} - \widehat{D}_j(z, a)\}^2 I(Z_{ij} = z)}{n_j - 1}, \\
\widehat{\xi}_{DE}^2(a) &= \frac{\sum_{i=1}^{J} \{\widehat{\mathsf{DED}}_j(a) - \widehat{\mathsf{DED}}(a)\}^2 I(A_j = a)}{J_a - 1}.
\end{aligned}
$$

First, for a given a, it is straightforward to show,

$$
(\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1} = \begin{pmatrix} \frac{\widehat{D}^2(1,a) + \widehat{D}^2(0,a)}{\widehat{\mathsf{DED}}^2(a)} & -\frac{\widehat{D}(1,a) + \widehat{D}(0,a)}{\widehat{\mathsf{DED}}^2(a)} \\ -\frac{\widehat{D}(1,a) + \widehat{D}(0,a)}{\widehat{\mathsf{DED}}^2(a)} & \frac{2}{\widehat{\mathsf{DED}}^2(a)} \end{pmatrix}.
$$

We then compute the projection matrix,

$$
\begin{aligned}
P_j &= \mathbf{W}_j^{1/2} \mathbf{M}_j (\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1} \mathbf{M}_j^\top \mathbf{W}_j^{1/2} \\
&= \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}} \mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(1, a) \mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}} \mathbf{1}_{n_{j0}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(0, a) \mathbf{1}_{n_{j0}} \end{pmatrix} \begin{pmatrix} \frac{\widehat{D}^2(1,a) + \widehat{D}^2(0,a)}{\widehat{\mathsf{DED}}^2(a)} & -\frac{\widehat{D}(1,a) + \widehat{D}(0,a)}{\widehat{\mathsf{DED}}^2(a)} \\ -\frac{\widehat{D}(1,a) + \widehat{D}(0,a)}{\widehat{\mathsf{DED}}^2(a)} & \frac{2}{\widehat{\mathsf{DED}}^2(a)} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}} \mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(1, a) \mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}} \mathbf{1}_{n_{j0}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(0, a) \mathbf{1}_{n_{j0}} \end{pmatrix}^\top \\
&= \begin{pmatrix} -\frac{1}{\sqrt{J_a n_{j1}}} \cdot \frac{\widehat{D}(0,a)}{\widehat{\mathsf{DED}}(1)} \mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \frac{1}{\widehat{\mathsf{DED}}(1)} \mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j1}}} \cdot \frac{\widehat{D}(1,a)}{\widehat{\mathsf{DED}}(1)} \mathbf{1}_{n_{j0}} & -\frac{1}{\sqrt{J_a n_{j1}}} \cdot \frac{1}{\widehat{\mathsf{DED}}(1)} \mathbf{1}_{n_{j0}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{J_a n_{j1}}} \mathbf{1}_{n_{j1}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(1, a) \mathbf{1}_{n_{j1}} \\ \frac{1}{\sqrt{J_a n_{j0}}} \mathbf{1}_{n_{j0}} & \frac{1}{\sqrt{J_a n_{j1}}} \cdot \widehat{D}(0, a) \mathbf{1}_{n_{j0}} \end{pmatrix}^\top \\
&= \begin{pmatrix} \frac{1}{J_a n_{j1}} \mathbf{1}_{n_{j1} \times n_{j1}} & \mathbf{0}_{n_{j1} \times n_{j0}} \\ \mathbf{0}_{n_{j0} \times n_{j1}} & \frac{1}{J_a n_{j0}} \mathbf{1}_{n_{j0} \times n_{j0}} \end{pmatrix}.
\end{aligned}
$$

Similar to the proof of Theorem 8, $(\mathbf{1}_{n_{j1}}^\top, \mathbf{0}_{n_{j0}}^\top)^\top$ and $(\mathbf{0}_{n_{j1}}^\top, \mathbf{1}_{n_{j0}}^\top)^\top$ are two eigenvectors of $I_{n_j} - P_b$ whose eigenvalue is $(J_a - 1)/J_a$. Thus, we have,

$$
(\boldsymbol{I}_{n_j} - \mathbf{P}_j)^{-1/2} \mathbf{W}_j \mathbf{M}_j = \sqrt{\frac{J_a}{J_a - 1}} \begin{pmatrix} \frac{1}{J_a n_{j1}} \mathbf{1}_{n_{j1}} & \frac{1}{J_a n_{j1}} \cdot \widehat{D}(1, a) \mathbf{1}_{n_{j1}} \\ \frac{1}{J_a n_{j0}} \mathbf{1}_{n_{j0}} & \frac{1}{J_a n_{j1}} \cdot \widehat{D}(0, a) \mathbf{1}_{n_{j0}} \end{pmatrix}.
$$

From the regression model given in equation (13), we obtain the following residuals for observations with $Z_{ij} = z$,

$$
\widehat{\eta}_{ij} = Y_{ij} - \widehat{Y}(z, a) - \widehat{\mathsf{CADE}}(a) \{D_{ij} - \widehat{D}(z, a)\}.
$$

This implies,

$$
\begin{aligned}
&\widehat{\boldsymbol{\eta}}_j^\top(\boldsymbol{I}_{n_j}-\mathbf{Q}_j)^{-1/2}\mathbf{W}\mathbf{M}_j\\
=&\sqrt{\frac{J_a}{J_a-1}}(Y_{1j}-\widehat{Y}(1,a)-\widehat{\mathsf{CADE}}(a)\{D_{ij}-\widehat{D}(1,a)\},\dots,Y_{n_jj}-\widehat{Y}(0,a)-\widehat{\mathsf{CADE}}(a)\{D_{ij}-\widehat{D}(0,a)\})\\
&\times\begin{pmatrix}\frac{1}{J_an_{j1}}\mathbf{1}_{n_{j1}} & \frac{1}{J_an_{j1}}\cdot\widehat{D}(1,a)\mathbf{1}_{n_{j1}}\\ \frac{1}{J_an_{j0}}\mathbf{1}_{n_{j0}} & \frac{1}{J_an_{j1}}\cdot\widehat{D}(0,a)\mathbf{1}_{n_{j0}}\end{pmatrix}\\
=&\sqrt{\frac{1}{J_a(J_a-1)}}\begin{pmatrix}\widehat{Y}_j(1,1)-\widehat{Y}(1,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(1,a)-\widehat{D}(1,a)\}\\ \widehat{D}(1,a)\{\widehat{Y}_j(1,a)-\widehat{Y}(1,a)\}+\widehat{D}(1,a)\widehat{\mathsf{CADE}}(1)\{\widehat{D}_j(1,1)-\widehat{D}(1,a)\}\end{pmatrix}^\top\\
&+\sqrt{\frac{1}{J_a(J_a-1)}}\begin{pmatrix}\widehat{Y}_j(0,a)-\widehat{Y}(0,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\\ \widehat{D}(0,a)\{\widehat{Y}_j(0,a)-\widehat{Y}(0,a)\}+\widehat{D}(0,a)\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\end{pmatrix}^\top
\end{aligned}
$$

Let $\mathbf{U}_j=\mathbf{M}_j^\top\mathbf{W}_j(\boldsymbol{I}_{n_j}-\mathbf{P}_j)^{-1/2}\widehat{\boldsymbol{\eta}}_j\widehat{\boldsymbol{\eta}}_j^\top(\boldsymbol{I}_{n_j}-\mathbf{P}_j)^{-1/2}\mathbf{W}_j\mathbf{M}_j$ and $u_{jk_1k_2}$ be the k_1k_2-th entry of $\mathbf{U}_j$,

$$
\begin{aligned}
u_{j11} &= \frac{1}{J_a(J_a-1)}\Big[\widehat{Y}_j(1,a)-\widehat{Y}(1,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(1,a)-\widehat{D}(1,a)\\
&\quad +\widehat{Y}_j(0,a)-\widehat{Y}(0,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\Big]^2,\\
u_{j12} &= \frac{1}{J_a(J_a-1)}\Big[\widehat{Y}_j(1,a)-\widehat{Y}(1,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(1,a)-\widehat{D}(1,a)\\
&\quad +\widehat{Y}_j(0,a)-\widehat{Y}(0,a)+\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\Big]\\
&\quad\times\Big[\widehat{D}(1,a)\{\widehat{Y}_j(1,a)-\widehat{Y}(1,a)\}+\widehat{D}(1,a)\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(1,a)-\widehat{D}(1,a)\}\\
&\quad +\widehat{D}(0,a)\{\widehat{Y}_j(0,a)-\widehat{Y}(0,a)\}+\widehat{D}(0,a)\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\Big],\\
u_{j22} &= \frac{1}{J_a(J_a-1)}\Big[\widehat{D}(1,a)\{\widehat{Y}_j(1,a)-\widehat{Y}(1,a)\}+\widehat{D}(1,a)\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(1,a)-\widehat{D}(1,a)\}\\
&\quad +\widehat{D}(0,a)\{\widehat{Y}_j(0,a)-\widehat{Y}(0,a)\}+\widehat{D}(0,a)\widehat{\mathsf{CADE}}(a)\{\widehat{D}_j(0,a)-\widehat{D}(0,a)\}\Big]^2.
\end{aligned}
$$

From the definition of the cluster-robust HC2 variance, we have

$$
\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}((\widehat{\beta}_a^{\text{w2sls}},\widehat{\beta}_{1a}^{\text{w2sls}})^\top) = (\mathbf{M}^\top\mathbf{W}\mathbf{M})^{-1}\left\{\sum_{j=1}^J\mathbf{U}_jI(A_j=a)\right\}(\mathbf{M}^\top\mathbf{W}\mathbf{M})^{-1}.
$$

We can then obtain that

$$
\begin{aligned}
&\widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\beta}_{1a}^{\text{2sls}})\\
=&\ \frac{1}{\widehat{\mathsf{DED}}^4(a)}\sum_j\Big[4u_{j22}-4\{\widehat{D}(1,a)+\widehat{D}(0,a)\}u_{j12}+\{\widehat{D}(1,a)+\widehat{D}(0,a)\}^2u_{j11}\Big]I(A_j=a)\\
=&\ \frac{1}{\widehat{\mathsf{DED}}^2(a)}\left\{\frac{\widehat{\sigma}^2_{DE}(a)}{J_a}-2\widehat{\mathsf{CADE}}(1)\frac{\widehat{\zeta}^2_{DE}(a)}{J_a}+\widehat{\mathsf{CADE}}^2(1)\frac{\widehat{\xi}^2_{DE}(a)}{J_a}\right\}.
\end{aligned}
$$

We then calculate the individual-robust generalization of HC2 variance,

$$\widehat{\text{var}}_{\text{hc2}}^{\text{ind}}((\widehat{\beta}_a^{\text{2sls}}, \widehat{\beta}_{1a}^{\text{2sls}})^\top) = (\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1} \left\{ \sum_{j=0}^{1} \sum_{i=1}^{n_j} w_{ij}^2 \widehat{\eta}_{ij}^{*2} (1 - Q_{ij})^{-1} \mathbf{M}_{ij} \mathbf{M}_{ij}^\top I(A_j = a) \right\} (\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1},$$

where Q_{ij} is the individual leverage and $\widehat{\eta}_{ij}^*$ is the adjusted residual to have $X_j^\top \widehat{\eta}_j^* = 0_{n_j}$,

$$\widehat{\eta}_{ij}^* = Y_{ij} - \widehat{Y}_j(z,a) - \widehat{\mathsf{CADE}}(a)\{D_{ij} - \widehat{D}_j(z,a)\}, \text{ if } Z_{ij} = z.$$

For units with $Z_{ij} = z$, we have,

$$Q_{ij} = \frac{w_{ij} J_a}{\widehat{\mathsf{DED}}^2(1)} \mathbf{M}_{ij}^\top \begin{pmatrix} \widehat{D}^2(1) + \widehat{D}^2(0) & -\widehat{D}(1) - \widehat{D}(0) \\ -\widehat{D}(1) - \widehat{D}(0) & 2 \end{pmatrix} \mathbf{M}_{ij} = \frac{1}{n_{jz}}.$$

Thus, we have

$$\begin{aligned} & \sum_{j=1}^{J} \sum_{i=1}^{n_j} w_{ij}^2 \widehat{\eta}_{ij}^{*2} I(A_j = a)(1 - Q_{ij})^{-1} \mathbf{M}_{ij} \mathbf{M}_{ij}^\top \\ = & \frac{1}{J_a^2} \sum_{j=1}^{J} I(A_j = a) \sum_{i=1}^{n_j} \mathbf{M}_{ij} \left\{ \frac{[Y_{ij} - \widehat{Y}_j(1,a) - \widehat{\mathsf{CADE}}(a)\{D_{ij} - \widehat{D}_j(1,a)\}]^2 I(Z_{ij} = 1)}{n_{j1}(n_{j1} - 1)} \right. \\ & \left. + \frac{[Y_{ij} - \widehat{Y}_j(0,a) - \widehat{\mathsf{CADE}}(a)\{D_{ij} - \widehat{D}_j(0,a)\}]^2 I(Z_{ij} = 0)}{n_{j0}(n_{j0} - 1)} \right\} \mathbf{M}_{ij}^\top \\ = & \frac{1}{J_a^2} \begin{pmatrix} S_1 + S_0 & \widehat{D}(1,a) S_1 + \widehat{D}(0,a) S_0 \\ \widehat{D}(1,a) S_1 + \widehat{D}(0,a) S_0 & \widehat{D}^2(1,a) S_1 + \widehat{D}^2(0,a) S_0 \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} S_1 &= \sum_{j=1}^{J} \left\{ \frac{\widehat{\sigma}_j^2(1,a) + \widehat{\mathsf{CADE}}(a) \widehat{\zeta}_j^2(1,a) + \widehat{\mathsf{CADE}}^2(a) \widehat{\xi}_j^2(1,a)}{n_{j1}} \right\} I(A_j = a), \\ S_0 &= \sum_{j=1}^{J} \left\{ \frac{\widehat{\sigma}_j^2(0,a) + \widehat{\mathsf{CADE}}(a) \widehat{\zeta}_j^2(0,a) + \widehat{\mathsf{CADE}}^2(a) \widehat{\xi}_j^2(0,a)}{n_j - n_{j1}} \right\} I(A_j = a). \end{aligned}$$

Putting all together,

$$\begin{aligned} & \widehat{\text{var}}_{\text{hc2}}^{\text{ind}}(\widehat{\beta}_{1a}^{\text{w2sls}}) \\ = & \frac{1}{J_a^2 \widehat{\mathsf{DED}}^4(a)} \sum_{j=1}^{J} \Big[4\widehat{D}^2(1,a) S_1 + \widehat{D}^2(0,a) S_0 - 4\{\widehat{D}(1,a) + \widehat{D}(0,a)\}\{\widehat{D}(1,a) S_1 + \widehat{D}(0,a) S_0\} \\ & + \{\widehat{D}(1,a) + \widehat{D}(0,a)\}^2 \{S_1 + S_0\} \Big] I(A_j = a) \\ = & \frac{1}{\widehat{\mathsf{DED}}^2(a)} \frac{S_1 + S_0}{J_a^2}. \end{aligned}$$

Therefore, we have,

$$\widehat{\text{var}}\left\{\widehat{\mathsf{CADE}}(a)\right\} = \left(1 - \frac{J_a}{J}\right) \widehat{\text{var}}_{\text{hc2}}^{\text{cluster}}(\widehat{\beta}_{1a}^{\text{w2sls}}) + \frac{J_a}{J} \widehat{\text{var}}_{\text{hc2}}^{\text{ind}}(\widehat{\beta}_{1a}^{\text{w2sls}}).$$

□

D Simulation Studies

We examine the performance of different variance estimators for the CADE. In particular, we compare our variance estimator with three other commonly used variance estimators: HC2 variance (MacKinnon and White, 1985), cluster-robust variance (Liang and Zeger, 1986), and cluster-robust HC2 variance (Bell and McCaffrey, 2002). Following the notation used in Section 3.5.2, the HC2 variance is defined as,

$$(\mathbf{M}^\top \mathbf{W}\mathbf{M})^{-1}\left\{\sum_{j=1}^{J}\sum_{i=1}^{n_j} w_{ij}^2 \widehat{\eta}_{ij}^2 (1-Q_{ij})^{-1}\mathbf{M}_{ij}\mathbf{M}_{ij}^\top\right\}(\mathbf{M}^\top \mathbf{W}\mathbf{M})^{-1},$$

whereas the cluster-robust variance is defined as,

$$(\mathbf{M}^\top \mathbf{W}\mathbf{M})^{-1}\left\{\sum_{j=1}^{J}\mathbf{M}_j^\top \mathbf{W}_j \widehat{\eta}_j \widehat{\eta}_j^\top \mathbf{W}_j \mathbf{M}_j\right\}(\mathbf{M}^\top \mathbf{W}\mathbf{M})^{-1}.$$

Finally, the cluster-robust HC2 variance is defined in equation (14).

Below, we consider three scenarios: no spillover effect of the treatment receipt on the outcome, no spillover effect of the treatment assignment on the treatment receipt, and both spillover effects present. In each scenario, we choose equal size n in all J clusters, and generate the data with three different settings regarding the number of clusters and cluster sizes while holding the total number of units $N = nJ$ constant: $(n = 10, J = 250)$, $(n = 250, J = 10)$, and $(n = 50, J = 50)$. We find that our proposed variance estimators perform well so long as the number of clusters is large. The HC2 variance estimator tends to underestimate the true variance while the cluster-robust HC2 variance estimator tends to overestimate it.

D.1 No Spillover Effect of Treatment Receipt on the Outcome

We first conduct a simulation study under the assumption of no spillover effect of treatment receipt on the outcome. In this scenario, Assumptions 1–6 are satisfied. We begin by defining the complier status variable $C_{ij}^*(a)$ for a given treatment assignment mechanism $a = 0, 1$ as,

$$C_{ij}^*(a) \;=\; \begin{cases} 0 & \text{if } D_{ij}(1,a) = D_{ij}(0,a) = 0, \\ 1 & \text{if } D_{ij}(1,a) = 1, D_{ij}(0,a) = 0, \\ 2 & \text{if } D_{ij}(1,a) = D_{ij}(0,a) = 1, \end{cases} \tag{A14}$$

where 0, 1, and 2 represent never-taker, complier, and always-taker, respectively. We sample $C_{ij}^*(a)$ from a categorical distribution with probabilities $(0.1, 0.6, 0.3)$ for $a = 1$ and $(0.3, 0.6, 0.1)$ for $a = 0$. We obtain the potential values $D_{ij}(z, a)$ from the realized value of $C_{ij}^*(a)$ according to equation (A14).

In the absence of the spillover effect of the treatment receipt on the outcome, the potential values of the outcome only depend on one's treatment receipt. Therefore, we first generate $Y_{ij}(D_{ij} = 0)$ via $Y_{ij}(D_{ij} = 0) \overset{\text{i.i.d.}}{\sim} N(0, 1)$ and then generate the $Y_{ij}(D_{ij} = 1)$ as,

$$\theta_j \;\overset{\text{i.i.d.}}{\sim}\; N(\theta, \sigma_b^2), \quad Y_{ij}(D_{ij} = 1) - Y_{ij}(D_{ij} = 0) \;\overset{\text{i.i.d.}}{\sim}\; N(\theta_j, \sigma_w^2),$$

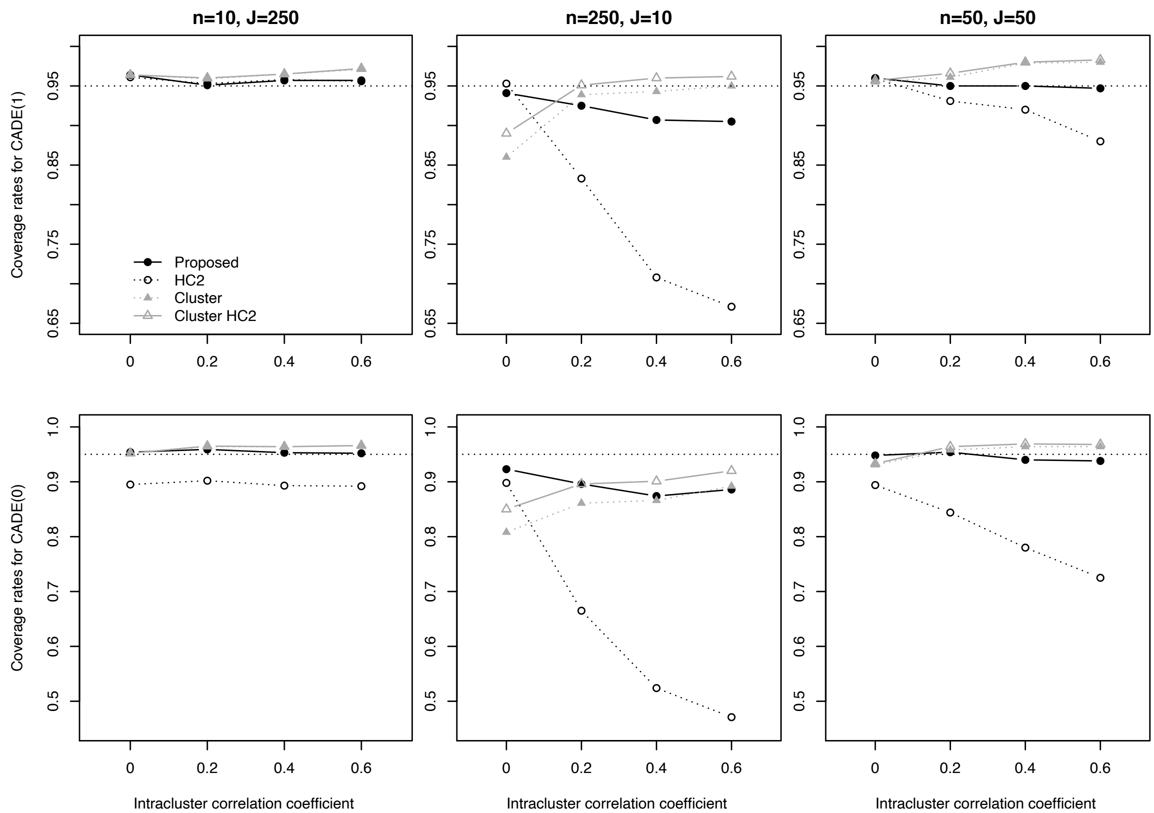

Figure A1: Coverage rates of 95% confidence intervals when there is no spillover effect of the treatment receipt on the outcome. The confidence intervals based on the proposed variance estimator (solid circle with black line) are compared with those based on the HC2 variance (open circle with dotted line), the cluster-robust variance (grey solid triangle with dotted line), and the cluster-robust HC2 variance (grey open triangle with solid line). The cluster size is indicated by n whereas J is the number of clusters. The horizontal axis represents the intracluster correlation coefficient.

where σ_b^2 represents the between-cluster variance and σ_w^2 is the within-cluster variance. We generate the treatment assignment mechanism A_j with $\Pr(A_j = a) = 1/2$ for $a = 0, 1$ such that $J_1 = J_0 = J/2$. We then completely randomize the treatment assignment Z_{ij} so that 60% (40%) of the units assigned to treatment if $A_j = 1$ ($A_j = 0$). For population parameters, we fix the average cluster specific effect $\theta = 1$ as well as the total variance $\sigma_b^2 + \sigma_w^2 = 1$. We use four different levels of intraclass correlation, i.e., $(0, 0.2, 0.4, 0.6)$, which is defined as $\rho = \sigma_b^2/(\sigma_b^2 + \sigma_w^2)$.

Figure A1 shows the coverage rates of the confidence intervals for the CADEs calculated by averaging over 1,000 Monte Carlo simulations (the top and bottom rows present the coverage rates for the $\mathsf{CADE}(1)$ and $\mathsf{CADE}(0)$, respectively). When the number of clusters is relatively large (i.e., $(n = 10, J = 250)$ and $(n = 50, J = 50)$), our variance estimator leads to the coverage rates closest to the nominal rate of 95%. However, when the number of cluster is small but the cluster size is large $(n = 250, J = 10)$, all the four variance estimators have a tendency to undercover the true value especially when the intracluster correlation is large. Across all scenarios, the confidence intervals based on the HC2 variance tend to perform poorly.

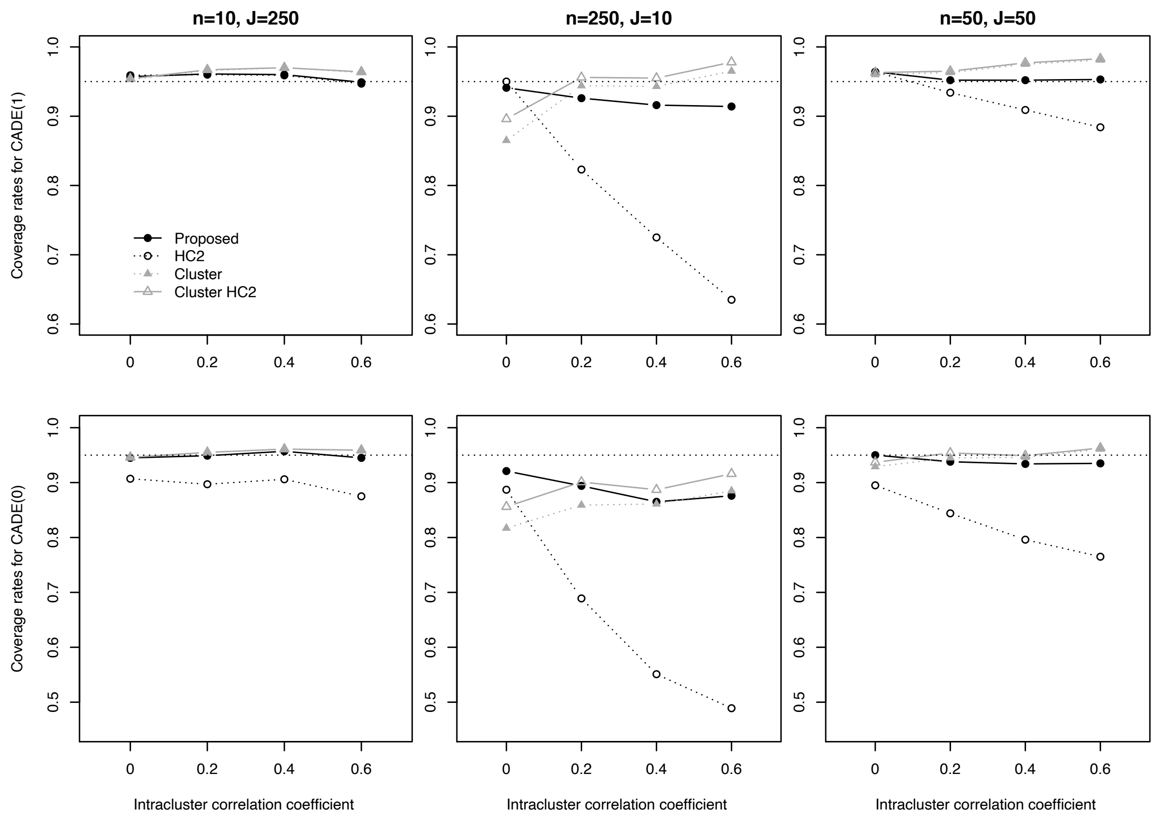

Figure A2: Coverage rates of 95% confidence intervals when there is no spillover effect of the treatment assignment on the treatment receipt. See the caption of Figure A1 for details.

D.2 No Spillover Effect of Treatment Assignment on the Treatment Receipt

We next consider the setting with no spillover effect of treatment assignment on the treatment receipt. In this scenario, Assumptions 1–5 are satisfied and Assumption 6 is violated. Since the potential values of treatment receipt depend only on one's own treatment assignment, the complier status does not depend on the treatment assignment mechanism, i.e., $C^*_{ij}(1) = C^*_{ij}(0) = C^*_{ij}$. We sample C^*_{ij} from a categorical distribution with probabilities $(0.2, 0.6, 0.2)$, and compute the potential values of Y_{ij} as,

$$Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j} = \boldsymbol{d}_{-i,j}) \overset{\text{indep.}}{\sim} N\left(\frac{\beta}{n}\sum_{i=1}^{n} d_{ij}, 1\right), \quad \theta_j \overset{\text{i.i.d.}}{\sim} N(\theta, \sigma_b^2),$$

$$Y_{ij}(D_{ij} = 1, \boldsymbol{D}_{-i,j} = \boldsymbol{d}_{-i,j}) - Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j} = \boldsymbol{d}_{-i,j}) \overset{\text{indep.}}{\sim} N(\theta_j, \sigma_w^2).$$

Thus, the potential values of the outcome depend on the number of treated units in the cluster. We then generate the treatment assignment and its mechanism in the same way as done in Section D.1 while setting $\theta = 1$ and $\beta = 1$. Figure A2 shows the coverage rates of the confidence intervals for the CADEs in the same way as in Figure A1. The results are similar to those under the setting with no spillover effect of the treatment receipt on the outcome shown in Figure A1.

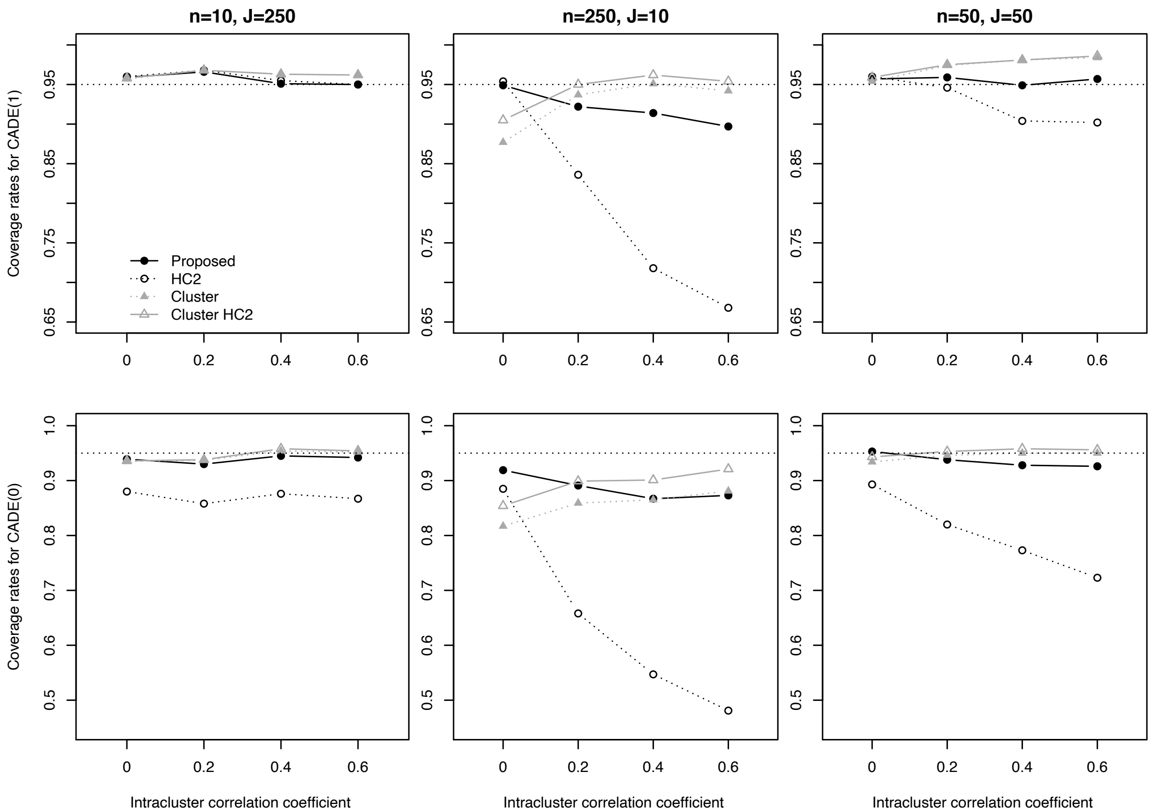

Figure A3: Coverage rates of 95% confidence intervals in the presence of two spillover effects. See the caption of Figure A1 for details.

D.3 Presence of Two Spillover Effects

Finally, we consider the setting with both types of spillover effects by combining the data generating mechanisms used above. We keep the data generating mechanism of the potential values of the treatment receipt introduced in Section D.1, while generating the outcome according to the data generating mechanism of Section D.2. Thus, this setting permits the presence of two spillover effects: the spillover effect of treatment receipt on the outcome and the spillover effect of treatment assignment on the treatment receipt. In this scenario, Assumptions 1–4 are satisfied and Assumptions 5 and 6 are violated. Figure A3 shows the coverage of the 95% confidence intervals for the CADEs. Again, the results are quite similar to those obtained under the other two scenarios.

D.4 Zero-Inflated Outcome

Because the outcome variable (annual hospital expenditure) in our application data has many zeros (20.7%) and is skewed, we consider a simulation study with a zero-inflated outcome variable. In particular, we maintain the core data generating mechanisms used in Sections D.1–D.3, and replace the distribution of the outcome variable with a mixture of the point mass at zero and a truncated normal distribution. For example, for the simulation setup used in Section D.1, we generate the outcome variable as,

$$U \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p),$$
$$Y_{ij}(D_{ij}=1) - Y_{ij}(D_{ij}=0) \stackrel{\text{i.i.d.}}{\sim} (1-U)\cdot \text{TruncN}(\theta_j, \sigma_w^2, (0, +\infty)),$$

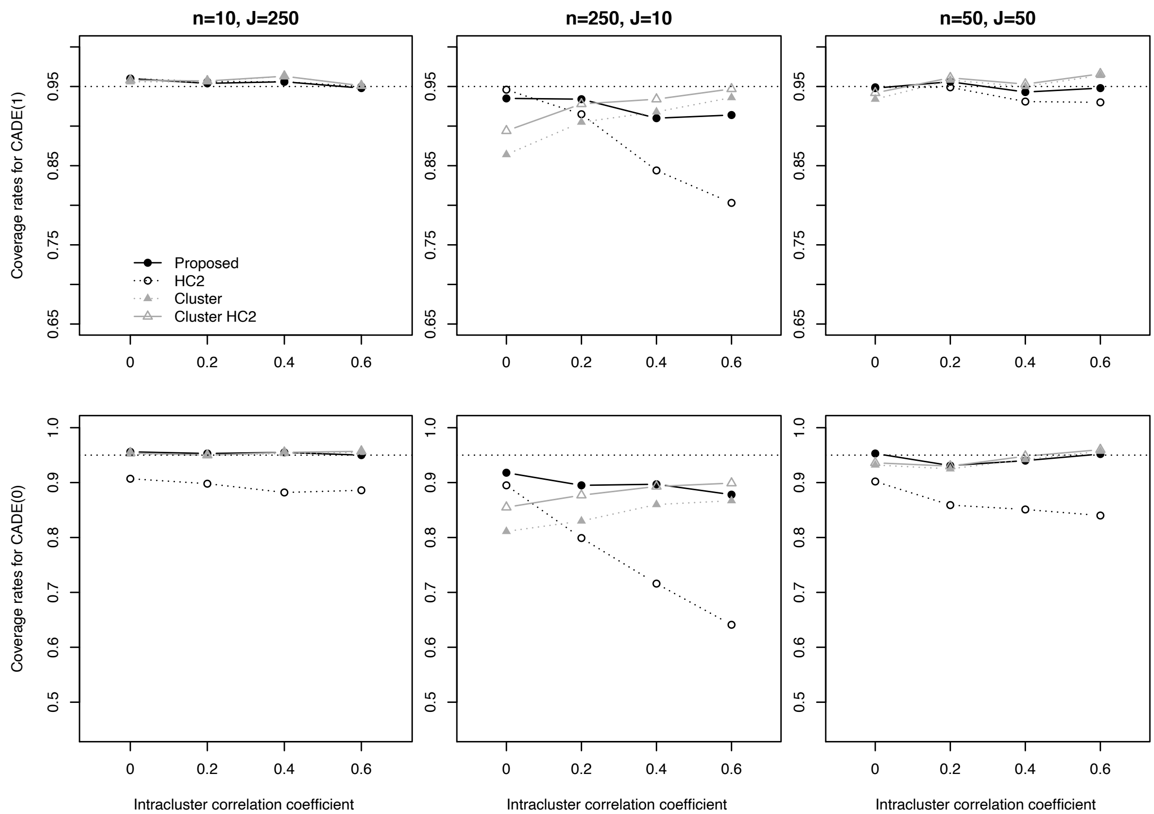

Figure A4: Coverage rates of 95% confidence intervals when there is no spillover effect o the treatment receipt on the outcome. See the caption of Figure A1 for details.

where $\text{TruncN}(\theta_j, \sigma_w^2, (0, +\infty))$ is a normal distribution with mean θ_j and variance σ_w^2 truncated from below at zero. We choose $p = 0.2$ to emulate our application data.

Figure A4 shows the coverage rates of the confidence intervals for the CADEs when the treatment receipt has no spillover effect on the outcome. The top and bottom rows present the coverage rates for the $\mathsf{CADE}(1)$ and $\mathsf{CADE}(0)$, respectively. The result is similar to that of the simulation in Section D.1. When the number of clusters is relatively large (i.e., $(n = 10, J = 250)$ and $(n = 50, J = 50)$), our variance estimator leads to the coverage rates closest to the nominal rate of 95%. However, when the number of cluster is small but the cluster size is large $(n = 250, J = 10)$, all the four variance estimators have a tendency to undercover the true value especially when the intracluster correlation is large.

Figure A5 presents the coverage of the 95% confidence intervals for the CADEs when there is no spillover effect of treatment assignment on the treatment receipt. Figure A6 shows the coverage of the 95% confidence intervals for the CADEs in the presence of two spillover effects. The results are similar to those with the corresponding cases of the normally distributed outcomes (see Sections D.2 and D.3).

E Model-Based Analysis

In the main text, we focus on the nonparametric identification of the ITT causal effects and complier causal effects and establish the connection between the randomization-based estimators and the regression-based estimators. Although the model-free identification results

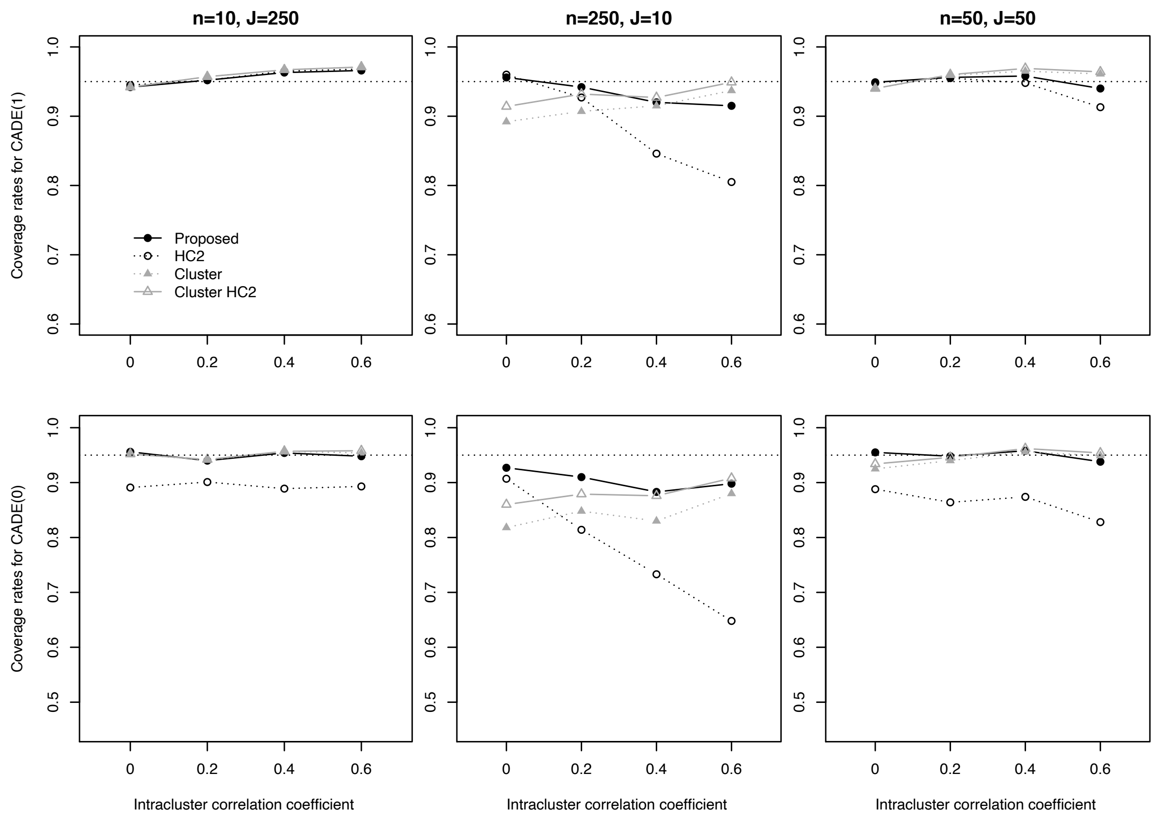

Figure A5: Coverage rates of 95% confidence intervals when there is no spillover effect of treatment assignment on the treatment receipt. See the caption of Figure A1 for details.

are appealing, they are limited by the experimental design and may not directly estimate the causal quantities that are of interest to applied researchers and policy makers. In this section, we consider a model-based analysis to overcome the limitations of our nonparametric approach. In the following, we assume linear models for the sake of illustration but other modeling assumptions are possible. We note that the model-based analysis is based on the super population framework, which is different from the finite population framework used for our nonparametric analysis.

E.1 Intention-to-Treat Analysis

We first consider the model-based ITT analysis. We model the potential outcome as a linear function of one's own encouragement, the proportion of encouraged households within the same village, and their interaction.

$$Y_{ij}(\boldsymbol{Z}_j = \boldsymbol{z}_j) \quad = \quad \alpha_0 + \alpha_1 z_{ij} + \alpha_2 \cdot \frac{\sum_{j=1}^{n_j} z_{ij}}{n_j} + \alpha_3 z_{ij} \cdot \frac{\sum_{j=1}^{n_j} z_{ij}}{n_j} + \iota_{ij}, \qquad \text{(A15)}$$

where ι_{ij} is the error term.

This model is applicable under two-stage randomized experiments with more than two treatment assignment mechanisms. In addition, the model can be used to extrapolate the average direct and spillover effects under different treatment assignment mechanisms. For example, under the scenario that all other households within the same cluster are assigned

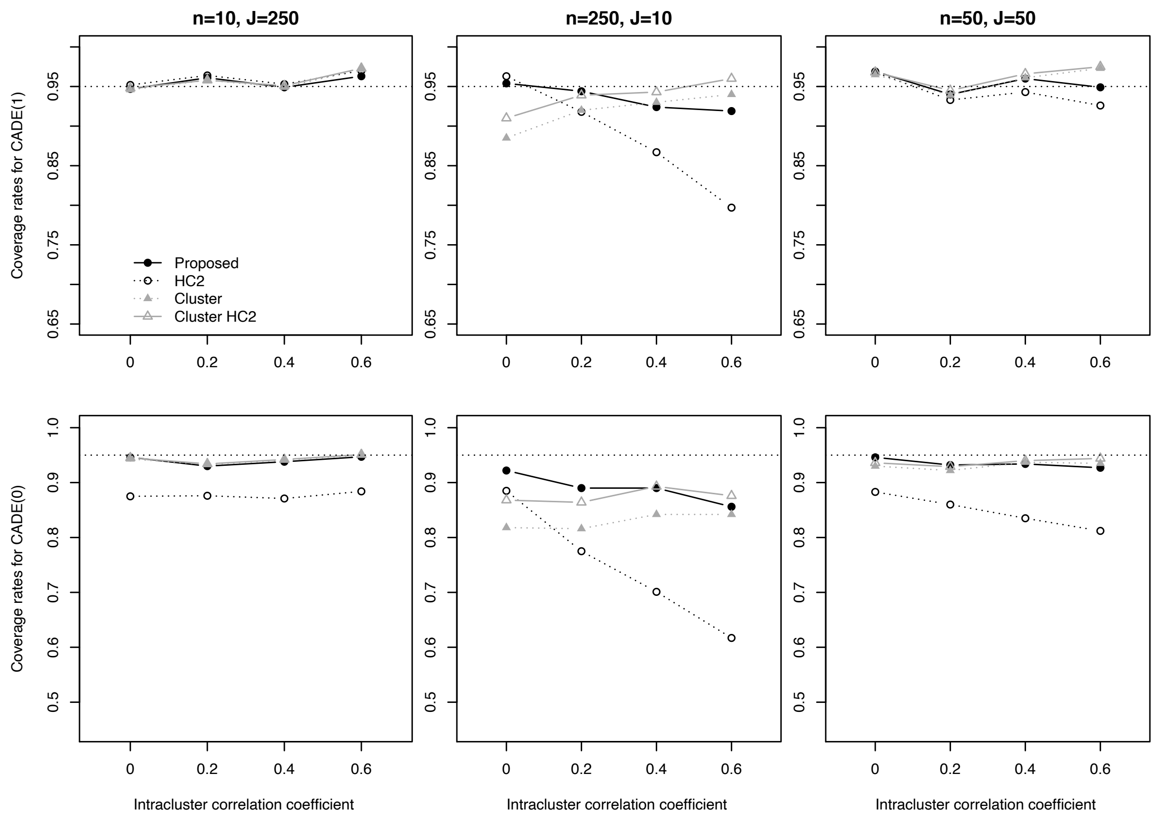

Figure A6: Coverage rates of 95% confidence intervals in the presence of two spillover effects. See the caption of Figure A1 for details.

to the treatment condition, the average direct effect is given by,

$$\mathbb{E}\{Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-i,j}=\mathbf{1})-Y_{ij}(Z_{ij}=0,\boldsymbol{Z}_{-i,j}=\mathbf{1})\} \approx \alpha_1+\alpha_3.$$

where the approximation results from an additional term α_2/n_j, which is negligible so long as n_j is large. If, on the other hand, all other households within the same cluster are assigned to the treatment condition, the average direct effect is equal to,

$$\mathbb{E}\{Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-i,j}=\mathbf{0})-Y_{ij}(Z_{ij}=0,\boldsymbol{Z}_{-i,j}=\mathbf{0})\} \approx \alpha_1.$$

Similarly, the average spillover effect of assigning all other households in the same cluster to the treatment condition (versus no household assigned to the treatment condition) depends on one's own encouragement status and is given by,

$$\begin{aligned}\mathbb{E}\{Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-i,j}=\mathbf{1})-Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-i,j}=\mathbf{0})\} &\approx \alpha_2+\alpha_3,\\ \mathbb{E}\{Y_{ij}(Z_{ij}=0,\boldsymbol{Z}_{-i,j}=\mathbf{1})-Y_{ij}(Z_{ij}=0,\boldsymbol{Z}_{-i,j}=\mathbf{0})\} &\approx \alpha_2.\end{aligned}$$

We apply this model-based approach to our application data. Table A1 shows the results. Although not statistically significant, the estimated average direct effect when all other households are assigned to the treatment condition ($\alpha_1+\alpha_3$) is negative whereas the estimated average direct effect when all other households are assigned to the control condition (α_1) is positive. The spillover effects are generally negative especially when a household is encouraged to enroll in the RSBY. These results are similar to those from the randomization-based approach.

Direct effects		Spillover effects	
$\alpha_1 + \alpha_3$	α_1	$\alpha_2 + \alpha_3$	α_2
-1253	1477	-2881	-180
$(-2646, 139)$	$(-94, 2988)$	$(-4631, -1131)$	$(-1991, 1630)$

Table A1: Estimated average direct and spillover effects under Model (A15). The first (second) column presents the estimated average direct effect when all other households within the same cluster are assigned to the treatment (control) condition. The third (fourth) column shows the estimated average spillover effect of all other households within the same cluster are assigned to the treatment condition versus no households are assigned to the treatment condition when the household itself is assigned to the treatment (control) condition. The confidence intervals are based on cluster-robust HC2 standard errors.

E.2 Complier Average Direct Effect of Treatment Receipt

Next, we consider a model-based approach to the estimation of the complier average direct effect (CADE). In the standard settings without interference between units, the complier average causal effect represents the average causal effect of one's own treatment receipt on the outcome among compliers (Angrist *et al.*, 1996). This interpretation is still applicable to our settings if there exists no spillover effect of encouragement on treatment receipt or no spillover effect of treatment receipt on outcome (i.e., Scenarios I and II of Figure 1). However, when both types of spillover effects exist (i.e., Scenario III of the figure), the CADE represents the indirect effect of one's own encouragement on the outcome through the treatment receipt of other units in the same village as well as the direct effect of one's own treatment receipt on the outcome. Unfortunately, without an additional assumption, we cannot identify the latter in the presence of these two spillover effects.

Here, we address this issue by assuming the following parametric structure for the spillover effects,

$$Y_{ij}(\boldsymbol{D}_j = \boldsymbol{d}_j) = \beta_0 + \beta_1 d_{ij} + \beta_2 \cdot \frac{\sum_{j=1}^{n_j} d_{ij}}{n_j} + \beta_3 d_{ij} \cdot \frac{\sum_{j=1}^{n_j} d_{ij}}{n_j} + U_{ij}, \tag{A16}$$

where U_{ij} represents the latent confounders between the treatment receipt and the outcome. Model (A16) posits the potential outcome as a linear function of one's own treatment receipt and the proportion of households in the same cluster who received the treatment.

Under this model, the average direct effect of one's own treatment receipt on the outcome when all the other households within the same cluster receive the treatment is given by,

$$\mathbb{E}\{Y_{ij}(D_{ij} = 1, \boldsymbol{D}_{-i,j} = \mathbf{1}) - Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j} = \mathbf{1})\} \approx \beta_1 + \beta_3.$$

Similarly, the average direct effect of one's own treatment receipt when all the other households receive the control condition is,

$$\mathbb{E}\{Y_{ij}(D_{ij} = 1, \boldsymbol{D}_{-i,j} = \mathbf{0}) - Y_{ij}(D_{ij} = 0, \boldsymbol{D}_{-i,j} = \mathbf{0})\} \approx \beta_1.$$

Similarly, the average spillover effects of assigning all households to the treatment condition (versus no households assigned to the treatment condition) depends on one's own

Direct effects		Spillover effects	
$\beta_1+\beta_3$	β_1	$\beta_2+\beta_3$	β_2
-6013	8724	-11715	3022
$(-11872,-154)$	$(407,17041)$	$(-19445,-3985)$	$(-4927,10970)$

Table A2: Estimated direct and spillover effects of the treatment receipt under Model (A16). The first (second) column presents the average direct effect of one's own treatment receipt when all other households within the same cluster receive the treatment (control) condition. The third (fourth) column shows the average spillover effect of all other households within the same cluster receive the treatment condition versus no households receive the treatment condition when the household itself receives the treatment (control) condition. The confidence intervals are based on cluster-robust HC2 standard errors.

treatment assignment status and is given by,

$$\begin{aligned}\mathbb{E}\{Y_{ij}(D_{ij}=1,\boldsymbol{D}_{-i,j}=\mathbf{1})-Y_{ij}(D_{ij}=1,\boldsymbol{D}_{-i,j}=\mathbf{0})\} &\approx \beta_2+\beta_3,\\ \mathbb{E}\{Y_{ij}(D_{ij}=0,\boldsymbol{D}_{-i,j}=\mathbf{1})-Y_{ij}(D_{ij}=0,\boldsymbol{D}_{-i,j}=\mathbf{0})\} &\approx \beta_2.\end{aligned}$$

The following theorem establishes the identification of these effects under Model (A16).

THEOREM A5 *Suppose Model (A16) is correctly specified and Assumption 1 holds. Then, the coefficients of the model are identified by solving the following estimating equations*

$$\sum_{i=1}^{J}\sum_{j=1}^{n_j}\left\{Y_{ij}-(\beta_0+\beta_1 D_{ij}+\beta_2\overline{D}_j+\beta_3 D_{ij}\overline{D}_j)\right\}\boldsymbol{H}_{ij}=\mathbf{0}_4,$$

where $\boldsymbol{H}_{ij}=(1,Z_{ij},A_j,Z_{ij}A_j)^\top$.

We fit this model to our application data. Table A2 shows the results. The estimated average direct effect of enrollment in the RSBY receipt when all other households are also enrolled ($\beta_1+\beta_3$) is negative while the estimated average direct effect when all other households are not enrolled (β_1) is positive. This finding is similar to the one obtained under our nonparametric approach.