Submitted to the Annals of Applied Statistics

SUPPLEMENTARY MATERIAL FOR "SPATIALLY RANDOMIZED DESIGNS CAN ENHANCE POLICY EVALUATION"

BY YING YANG[1,a], CHENGSHUN SHI[2,b] FANG YAO[3,c]
SHOUYANG WANG[4,d] AND HONGTU ZHU[5,e]

[1]*Center for Applied Mathematics, Shanghai Key Laboratory for Contemporary Applied Mathematics, Fudan University,* [a]*yangying@fudan.edu.cn*

[2]*London School of Economics and Political Science,* [b]*C.Shi7@lse.ac.uk*

[3]*School of Mathematical Sciences, Center for Statistical Science, Peking University,* [c]*fyao@math.pku.edu.cn*

[4]*Academy of Mathematics and Systems Science, Chinese Academy of Sciences,* [d]*sywang@amss.ac.cn*

[5]*Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill,* [e]*htzhu@email.unc.edu*

Section 1 of this supplementary materials contains the additional implementation details (propensity score estimation in DR estimate and μ in DRL estimate). In Section 2, we present the supplementary theoretical results that complements the conclusions in the main text. The proofs for main theoretical conclusions are collected in Section 3. In Section 4, we introduce more noise structures and present results of the nondynamic settings under these noises to further verify the robustness of our conclusion.

1. Additional Implementation Details.

1.1. *Propensity score estimation in DR estimate.* Within the framework of spatially randomized designs, π_ι is predetermined and remains unaffected by observational data. Specifically, under the individual-randomized and cluster-randomized designs, the propensity scores are denoted by π_ι^i and π_ι^c, respectively. For the individual-randomized design, the propensity score for any unit ι and its neighbors can be expressed as $\pi_\iota^i(a) = p_\iota^a(1-p_\iota)^{1-a}\prod_{j\in\mathcal{N}_\iota}[p_j^a(1-p_j)^{1-a}]$. In the context of the cluster-randomized design, for units ι in a cluster $\mathcal{C}_j^0$, the propensity score simplifies to $\pi_\iota^c(a) = (p^{(j)})^a(1-p^{(j)})^{1-a}$. Conversely, for units ι on the boundary of a cluster $\partial\mathcal{C}_j$, the propensity score is given by $\pi_\iota^c(a) = (p^{(j)})^a(1-p^{(j)})^{1-a}\prod_{k\neq j}\left[\left\{1-\prod_{i\in\mathcal{N}_\iota}\mathbb{I}(i\notin\mathcal{C}_k)\right\}(p^{(k)})^a(1-p^{(k)})^{1-a}+\prod_{i\in\mathcal{N}_\iota}\mathbb{I}(i\notin\mathcal{C}_k)\right]$. This formulation accounts for the complex interplay between the treatment assignments of interference neighbors and the spatial structure of the clusters.

In the global design framework, we generate a sequence of i.i.d. Bernoulli(p) random variables denoted by $\{A_t^g\}_t$. For any given time point t, we set $A_{i,1} = \cdots = A_{i,R} = A_t^g$. This uniform treatment assignment simplifies the propensity score function $\pi_\iota^g(a|\{O_{i,j}\}_j)$ to $p^a(1-p)^{1-a}$, reflecting the global design's homogenous treatment distribution. To estimate ATE under this design, denoted as $\widehat{\tau}_{DR}^g$, we again leverage the same cross-fitting technique.

1.2. *Estimation of μ in DRL estimate..* We introduce the estimation method in this section. Let $\widetilde{O}_{\iota t} = (O_{\iota t}, |\mathcal{N}_\iota|^{-1}\sum_{k\in\mathcal{N}_\iota} O_{kt})$ and $\bar{m}_\iota(\boldsymbol{A}_t) = |\mathcal{N}_\iota|^{-1}\sum_{k\in\mathcal{N}_\iota} A_{kt}$. Recalling the expression of $\mu_{\iota t}^a$, it is sufficient to estimate $\omega_\iota^a(\widetilde{O}_{\iota t}) = p_a(\widetilde{O}_{\iota t})/p_b(\widetilde{O}_{\iota t})$. When dealing with stationary observations, $\omega_{\iota t}$ can be estimated

$$\widehat{\omega}_\iota^a = \arg\min_{\omega_\iota\in\Omega}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{M-1}\Delta_{\iota t}^a(\omega_\iota)f(\widetilde{O}_{\iota,t+1})\right|^2,$$

where Ω and $\mathcal{F}$ are some function classes, and

$$\Delta_{\iota t}^{a}(\omega_\iota) = \omega_\iota(\widetilde{O}_{\iota,t}) \frac{\mathbb{I}\{A_{\iota t} = a, \bar{m}_\iota(\boldsymbol{A}_t) = \bar{m}_\iota(a)\}}{P(A_{\iota t} = a, \bar{m}_\iota(\boldsymbol{A}_t) = \bar{m}_\iota(a) | \widetilde{O}_{\iota t})} - \omega_\iota(\widetilde{O}_{\iota,t+1}).$$

By further specifying Ω and $\mathcal{F}$ as particular function classes, it's possible to obtain a closed-form expression for the objective function. This allows for solving the optimization problem using stochastic gradient descent algorithms or neural networks. As elaborated in Shi et al. (2022), for high-dimensional $O_{\iota t}$, Ω and $\mathcal{F}$ can be designated as classes of deep neural networks, while for simpler scenarios, linear functions suffice. We model ω_ι as a linear blend of random Fourier basis functions $\{b_i(\widetilde{O})\}_{t=1}^{\infty}$, and choose $\mathcal{F}$ as the unit ball in a reproducing kernel Hilbert space (RKHS) characterized by the kernel $\kappa(\widetilde{O}_1, \widetilde{O}_2)$. Assume $\omega_\iota(\widetilde{O})$ is represented as $\sum_{t=1}^{d_\omega} \beta_{\omega,\iota,i} b_i(\widetilde{O})$, simplified to $\beta_{\omega,\iota}^\top b_\omega(\widetilde{O})$, where $\|\beta_{\omega,\iota}\|_2 = (\sum_{t=1}^{d_\omega} \beta_{\omega,\iota,i}^2)^{1/2} = 1$. By Lemma 1 of Shi et al. (2023), $\beta_{\iota,\omega}$ is the solution to

$$\arg\min_{\|\beta\|_2=1} \sup_{f\in\mathcal{F}} \left| \sum_{t=1}^{M-1} \beta^\top \delta_{\iota t} f(\widetilde{O}_{\iota,t+1}) \right|^2 = \arg\min_{\|\beta\|_2=1} \mathbb{E} \sum_{t_1=1}^{M-1} \sum_{t_2=1}^{M-1} \beta^\top \delta_{\iota t_1} \delta_{\iota t_2}^\top \beta \kappa(\widetilde{O}_{\iota,t_1+1}, \widetilde{O}_{\iota,t_2+1}),$$

where

$$\delta_{\iota t} = b_\omega(\widetilde{O}_{\iota t}) \frac{\mathbb{I}\{A_{\iota t} = a, \bar{m}_\iota(\boldsymbol{A}_t) = \bar{m}_\iota(a)\}}{P(A_{\iota t} = a, \bar{m}_\iota(\boldsymbol{A}_t) = \bar{m}_\iota(a) | \widetilde{O}_{\iota t})} - b_\omega(\widetilde{O}_{\iota,t+1}).$$

Then, we can estimate $\beta_{\iota,\omega}$ by

$$\widehat{\beta}_{\iota,\omega} = \arg\min_{\|\beta\|_2=1} \beta^\top \left\{ \sum_{i=1}^{N} \sum_{t_1=1}^{M-1} \sum_{t_2=1}^{M-1} \delta_{i\iota t_1} \delta_{i\iota t_2}^\top \kappa(\widetilde{O}_{i,\iota,t_1+1}, \widetilde{O}_{i,\iota,t_2+1}) \right\} \beta.$$

2. Supplementary Theoretical Foundations. In this section, we present the supplementary theoretical results that complements the conclusions in the main text. All proofs are deferred to the next section.

2.1. *Estimation accuracy in the nondynamic setting.* We write $q = p(1-p)$, $q_\iota = p_\iota(1-p_\iota)$ for $\iota = 1, \dots, R$ and $q^{(j)} = p^{(j)}(1-p^{(j)})$ for $j = 1, \cdots, m$. For the parametric model, we have the following conclusion for arbitary p_ι and $p^{(j)}$.

THEOREM S.1. *Suppose that CA holds. Recall that n_ι is the number of interference neighbors of the ιth region. Let $n_\iota^{(j)}$ denote the number of interference neighbors of to the ιth region belonging to the jth cluster. Then as $N \to \infty$, we have*

$$\mathrm{MSE}\,(\widehat{\tau}^g) \asymp \sum_{\iota=1}^{R} \sum_{\iota'=1}^{R} \frac{\mathbb{V}_{\iota\iota'}}{Nq},$$

$$\mathrm{MSE}\,(\widehat{\tau}^i) \asymp \frac{1}{N} \sum_{\iota=1}^{R} \frac{1}{q_\iota} \mathbb{V}_{\iota\iota} + \frac{2}{N} \sum_{\iota=1}^{R} \sum_{\iota'\in\mathcal{N}_\iota} \frac{n_\iota}{\sum_{k\in\mathcal{N}_\iota} q_k} \mathbb{V}_{\iota,\iota'} + \frac{1}{N} \sum_{\iota,\iota'=1}^{R} \frac{n_\iota n_{\iota'} \sum_{k\in\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}} q_k}{\sum_{k_1\in\mathcal{N}_\iota} q_{k_1} \sum_{k_2\in\mathcal{N}_{\iota'}} q_{k_2}} \mathbb{V}_{\iota\iota'},$$

$$\begin{aligned}\mathrm{MSE}(\widehat{\tau}^c) \asymp & \frac{1}{N} \sum_{j=1}^{m} \frac{\sum_{\iota,\iota'\in\mathcal{C}_j} \mathbb{V}_{\iota\iota'}}{q^{(j)}} + \frac{2}{N} \sum_{j=1}^{m} \sum_{j'\neq j} \sum_{\iota\in\mathcal{C}_j} \sum_{\iota'\in\partial\mathcal{C}_{j'}} \frac{\sum_{k_1\neq j'} n_{\iota'}^{(k_1)} n_{\iota'}^{(j)}}{\sum_{k\neq j'} (n_{\iota'}^{(k)})^2 q^{(k)}} \mathbb{V}_{\iota\iota'} \\ & + \frac{1}{N} \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{\iota\in\partial\mathcal{C}_{j_1}} \sum_{\iota'\in\partial\mathcal{C}_{j_2}} \frac{\sum_{k\neq j_1,j_2} n_\iota^{(k)} n_{\iota'}^{(k)} q^{(k)} \sum_{k\neq j_1} n_\iota^{(k)} \sum_{k\neq j_2} n_{\iota'}^{(k)}}{\sum_{k\neq j_1} (n_\iota^{(k)})^2 q^{(k)} \sum_{k\neq j_2} (n_{\iota'}^{(k)})^2 q^{(k)}} \mathbb{V}_{\iota\iota'},\end{aligned}$$

where $a_N \asymp b_N$ means that $a_N / b_N \to 1$ as $N \to \infty$.

For the nonparametric learning, the following theorem holds for $p_\iota = p^{(j)} = p$ where $0 < p < 1$.

THEOREM S.2. *Suppose that CA holds and $p_\iota = p^{(j)} = p$ for any ι and j. Then, we have*

$$\begin{aligned}
\mathrm{MSE}(\widehat{\tau}^g_{DR}) &= \frac{1}{N}\sigma_O^2 + \frac{1}{Np(1-p)}\sum_{\iota=1}^R\sum_{\iota'=1}^R \mathbb{V}_{\iota\iota'} + O\Big[\frac{R}{Np(1-p)}\sum_{\iota=1}^R \delta^2_{N,\iota}\Big],\\
\mathrm{MSE}(\widehat{\tau}^i_{DR}) &= \frac{1}{N}\sigma_O^2 + \frac{1}{N}\sum_{\iota}\sum_{\iota'}\Big(\frac{1}{p^{n_{\iota\iota'}}} + \frac{1}{(1-p)^{n_{\iota\iota'}}}\Big)\mathbb{V}_{\iota\iota'}\mathbb{I}(n_{\iota\iota'}>0)\\
&\quad + O\Big[\frac{r^2}{N}\Big(\frac{1}{p^{r+1}} + \frac{1}{(1-p)^{r+1}}\Big)\sum_{\iota=1}^R \delta^2_{N,\iota}\Big],\\
\mathrm{MSE}(\widehat{\tau}^c_{DR}) &= \frac{1}{N}\sigma_O^2 + \frac{1}{N}\sum_{\iota}\sum_{\iota'}\Big(\frac{1}{p^{m_{\iota\iota'}}} + \frac{1}{(1-p)^{m_{\iota\iota'}}}\Big)\mathbb{V}_{\iota\iota'}\mathbb{I}(m_{\iota\iota'}>0)\\
&\quad + O\Big[\frac{1}{Np(1-p)}\sum_{j=1}^m |\mathcal{C}_j^0|\sum_{\iota\in\mathcal{C}_j^0}\delta^2_{N,\iota}\Big]\\
&\quad + O\Big[\frac{m'}{N}\Big(\frac{1}{p^{r'+1}} + \frac{1}{(1-p)^{r'+1}}\Big)\sum_{j=1}^m |\partial\mathcal{C}_j|\sum_{\iota\in\partial\mathcal{C}_j}\delta^2_{N,\iota}\Big].
\end{aligned}$$

where $n_{\iota\iota'} = |(\mathcal{N}_\iota\cup\{\iota\})\cap(\mathcal{N}_{\iota'}\cup\{\iota'\})|$, $m_{\iota\iota'} = \sum_{k=1}^m \mathbb{I}((\mathcal{N}_\iota\cup\{\iota\})\cap\mathcal{C}_k\neq\emptyset, (\mathcal{N}_{\iota'}\cup\{\iota'\})\cap\mathcal{C}_k\neq\emptyset)$, $m' = \max_k |\{k' : \mathcal{N}_{\mathcal{C}_k}\cap\mathcal{N}_{\mathcal{C}_{k'}}\neq\emptyset\}|$.

The above MSEs are primarily determined by the first two terms. The residual term, governed by the rate at which the estimated outcome regression function converges, decays to zero at a faster rate. Notably, the first term is shared across the MSEs of all three estimators, which is independent of the treatment. The second term introduces a divergence in the MSEs of these estimators, which we will explore in detail in the following corollary.

2.2. *Estimation accuracy and inference in the dynamic setting.* We first present the MSEs of the ATE estimates of the parametric learning.

THEOREM S.3. *Assume that either the constant design, the independent design, or the switchback design is implemented temporally. As $N\to\infty$, we have:*

$$\mathrm{MSE}\left(\widehat{\tau}^g_{OLS}\right) - N^{-1}\sigma^2_{OLS} \asymp \frac{4}{N}\sum_{t=1}^M\sum_{\iota=1}^R\sum_{\iota'=1}^R \mathbb{V}^u_{\iota\iota' t},$$

$$\mathrm{MSE}\left(\widehat{\tau}^i_{OLS}\right) - N^{-1}\sigma^2_{OLS} \asymp \frac{4}{N}\sum_{t=1}^M\left\{\sum_{\iota=1}^R \mathbb{V}^u_{\iota\iota t} + 2\sum_{\iota=1}^R\sum_{\iota'\in\mathcal{N}_\iota}\mathbb{V}^u_{\iota\iota' t} + \sum_{\iota,\iota'=1}^R |\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}|\,\mathbb{V}^u_{\iota\iota' t}\right\},$$

$$\begin{aligned}
\mathrm{MSE}(\widehat{\tau}^c_{OLS}) - N^{-1}\sigma^2_{OLS} \asymp \frac{4}{N}\sum_{t=1}^M\sum_{j=1}^m\Bigg\{&\sum_{\iota,\iota'\in\mathcal{C}_j}\mathbb{V}^u_{\iota\iota' t} + 2\sum_{\iota\in\mathcal{C}_j}\sum_{\iota'\in\mathcal{N}_{\mathcal{C}_j}}\omega_{\iota'} n^{(j)}_{\iota'}\mathbb{V}^u_{\iota\iota' t} +\\
&\sum_{\iota,\iota'\in\mathcal{N}_{\mathcal{C}_j}}\omega_\iota\omega_{\iota'} n^{(j)}_\iota n^{(j)}_{\iota'}\mathbb{V}^u_{\iota\iota' t}\Bigg\},
\end{aligned}$$

where σ^2_{OLS} represents the common part of the MSEs of the three estimators and $\omega_\iota = \sum_{k\neq j} n_\iota^{(k)} / \sum_{k\neq j}(n_\iota^{(k)})^2$ for any ι.

From Theorem S.3, it is straightforward to establish Theorem 3.2. Now we present the asymptotic normality of $\widehat{\tau}_{OLS}$ and $\widehat{\tau}_{DRL}$ when M remains finite.

THEOREM S.4. *For models and ATE estimators discussed in this section, we have the following asymptotic normality results:*

(6.1) *when conditions of Theorem S.3 hold, there exist constants $v^g_{OLS}, v^i_{OLS}, v^c_{OLS}$ such that as $N\to\infty$, $\sqrt{N(v^g_{OLS})^{-1}}\left(\widehat{\tau}^g_{OLS}-\tau\right)$, $\sqrt{N(v^i_{OLS})^{-1}}\left(\widehat{\tau}^i_{OLS}-\tau\right)$ and $\sqrt{N(v^c_{OLS})^{-1}}\left(\widehat{\tau}^c_{OLS}-\tau\right)$ are asymtotically standard normal distributed.*

(6.2) *when the assumptions of Theorem 3.4 hold, there exist constants $v^g_{DRL}, v^i_{DRL}, v^c_{DRL}$ such that as $N\to\infty$, $\sqrt{N(v^g_{DRL})^{-1}}\left(\widehat{\tau}^g_{DRL}-\tau\right)$, $\sqrt{N(v^i_{DRL})^{-1}}\left(\widehat{\tau}^i_{DRL}-\tau\right)$ and $\sqrt{N(v^c_{DRL})^{-1}}\left(\widehat{\tau}^c_{DRL}-\tau\right)$ are asymtotically standard normal distributed.*

The first conclusion is derived from the asymptotic normality of OLS estimators, while the second is inferred from Theorem 19 in Shi et al. (2023). Consistent variance estimators can be acquired utilizing the bootstrap method. Similar to Section 2.4, with the asymptotic normality and consistent variance estimators, we can establish the Walt test statistics. Recalling that in Theorem S.3 and 3.4, we have proved that the spatially randomized designs can produce estimators with smaller MSEs. Note that the testing power decreases with MSEs. We can deduce that spatially randomized designs can lead to testing statistics with higher powers.

3. Proofs for the Main Theorems.

3.1. *Proof of Theorem S.1.* We first consider the global design. Let M^g_ι denote the corresponding design matrix

$$\begin{pmatrix} 1 & A_{1\iota} & O_{1\iota}^\top \\ 1 & A_{2\iota} & O_{2\iota}^\top \\ \vdots & \vdots & \vdots \\ 1 & A_{N\iota} & O_{N\iota}^\top \end{pmatrix}.$$

The OLS estimators are given by

$$\begin{pmatrix} \widehat{\alpha}_\iota \\ \widehat{\gamma}^g_\iota \\ \widehat{\beta}_\iota \end{pmatrix} = \begin{pmatrix} \alpha_\iota \\ \gamma^g_\iota \\ \beta_\iota \end{pmatrix} + [(M^g_\iota)^\top M^g_\iota]^{-1}(M^g_\iota)^\top e_\iota, \tag{1}$$

where $e_\iota = (e_{1\iota},\ldots,e_{N\iota})^\top \in \mathbb{R}^{N\times 1}$. Under the given conditions, it is immediate to see that

$$\frac{1}{N}(M^g_\iota)^\top M^g_\iota = \begin{pmatrix} 1 & \mathbb{E}A_{i\iota} & \mathbb{E}O_{i\iota}^\top \\ \mathbb{E}A_{i\iota} & \mathbb{E}A_{i\iota}^2 & \mathbb{E}A_{i\iota}O_{i\iota}^\top \\ \mathbb{E}O_{i\iota} & \mathbb{E}A_{i\iota}O_{i\iota} & \mathbb{E}O_{i\iota}O_{i\iota}^\top \end{pmatrix} + O_p\left(N^{-1/2}\right). \tag{2}$$

Since in the global randomized design, $A_{i\iota} = A_i \overset{\text{i.i.d}}{\sim} \text{Bernoulli}(p_\iota)$ with $p = 0.5$, independent of $\{(e_{i\iota j}, O_{i\iota j}^\top)\}_{i,j,\iota}$, we have that

$$\mathbb{E}A_{i\iota} = \mathbb{E}A_{i\iota}^2 = p,\ \mathbb{E}A_{i\iota}O_{i\iota} = p\mathbb{E}O_{i\iota}.$$

Notice that

$$\begin{pmatrix} 1 & 0 & 0^\top \\ -p & 1 & 0^\top \\ -\mathbb{E}O_{i\iota} & 0 & I \end{pmatrix} \begin{pmatrix} 1 & p & \mathbb{E}O_{i\iota}^\top \\ p & p & p\mathbb{E}O_{i\iota}^\top \\ \mathbb{E}O_{i\iota} & p\mathbb{E}O_{i\iota} & \mathbb{E}O_{i\iota}O_{i\iota}^\top \end{pmatrix} \begin{pmatrix} 1 & -p & -\mathbb{E}O_{i\iota} \\ 0 & 1 & 0 \\ 0 & 0^\top & I \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0^\top \\ 0 & p(1-p) & 0^\top \\ 0 & 0 & \mathrm{cov}(O_{i\iota}) \end{pmatrix}.$$

It follows that

$$\begin{aligned} \begin{pmatrix} 1 & p & \mathbb{E}O_{i\iota}^\top \\ p & p & p\mathbb{E}O_{i\iota}^\top \\ \mathbb{E}O_{i\iota} & p\mathbb{E}O_{i\iota} & \mathbb{E}O_{i\iota}O_{i\iota}^\top \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & -p & -\mathbb{E}O_{i\iota} \\ 0 & 1 & 0 \\ 0 & 0^\top & I \end{pmatrix} \begin{pmatrix} 1 & 0 & 0^\top \\ 0 & \dfrac{1}{p(1-p)} & 0^\top \\ 0 & 0 & \mathrm{cov}^{-1}(O_{i\iota}) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0^\top \\ -p & 1 & 0^\top \\ -\mathbb{E}O_{i\iota} & 0 & I \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{1}{1-p} + (\mathbb{E}O_{i\iota})^\top \mathrm{cov}^{-1}(O_{i\iota})\mathbb{E}O_{i\iota} & -\dfrac{1}{1-p} & -(\mathbb{E}O_{i\iota})^\top \mathrm{cov}^{-1}(O_{i\iota}) \\ -\dfrac{1}{1-p} & \dfrac{1}{p(1-p)} & 0^\top \\ -\mathrm{cov}^{-1}(O_{i\iota})\mathbb{E}O_{i\iota} & 0 & \mathrm{cov}^{-1}(O_{i\iota}) \end{pmatrix}. \end{aligned}$$

As such, we obtain that

$$\begin{aligned} \left\{\frac{1}{N}(M_\iota^g)^\top M_\iota^g\right\}^{-1} = &\begin{pmatrix} \dfrac{1}{1-p} + (\mathbb{E}O_{i\iota})^\top \mathrm{cov}^{-1}(O_{i\iota})\mathbb{E}O_{i\iota} & -\dfrac{1}{1-p} & -(\mathbb{E}O_{i\iota})^\top \mathrm{cov}^{-1}(O_{i\iota}) \\ -\dfrac{1}{1-p} & \dfrac{1}{p(1-p)} & 0^\top \\ -\mathrm{cov}^{-1}(O_{i\iota})\mathbb{E}O_{i\iota} & 0 & \mathrm{cov}^{-1}(O_{i\iota}) \end{pmatrix} \\ &+O_p\left(N^{-1/2}\right). \end{aligned}$$

Substituting it into (1), we have

$$\widehat{\tau}^g = \tau^g + \frac{1+o_p(1)}{N}\sum_{i=1}^N \sum_{\iota=1}^R \frac{A_{i\iota}-p}{p(1-p)} e_{i\iota}.$$

Its MSE is thus given by

$$\begin{aligned} \mathrm{MSE}(\widehat{\tau}^g) &= \frac{1+o(1)}{Np^2(1-p)^2}\mathrm{Var}\left\{\sum_{\iota=1}^R (A_{i\iota}-p)e_{i\iota}\right\} \\ &= \frac{1+o(1)}{Np^2(1-p)^2}\sum_{\iota,\iota'} \mathrm{cov}\{(A_i-p)e_{i\iota},(A_i-p)e_{i\iota'}\} \\ &= \frac{1+o(1)}{Np(1-p)}\sum_{\iota,\iota'}\mathbb{V}_{\iota\iota'} = \frac{4+o(1)}{N}\sum_{\iota,\iota'}\mathbb{V}_{\iota\iota'}. \end{aligned}$$

Next, we consider a special cluster-randomized design, corresponding to the individual-randomized design. Under this design, $\{A_{i\iota}\}_{i\iota} \overset{\text{i.i.d}}{\sim} \text{Bernoulli}(p)$. Let M_ι^i denote

$$\begin{pmatrix} 1 & A_{1\iota} & \overline{A}_{1\mathcal{N}_\iota} & O_{1\iota}^\top \\ 1 & A_{2\iota} & \overline{A}_{2\mathcal{N}_\iota} & O_{2\iota}^\top \\ \vdots & \vdots & \vdots & \vdots \\ 1 & A_{N\iota} & \overline{A}_{N\mathcal{N}_\iota} & O_{N\iota}^\top \end{pmatrix}.$$

Similar to (2), we have

$$\frac{1}{N}(M_\iota^i)^\top M_\iota^i=\begin{pmatrix} 1 & \mathbb{E}A_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}O_{i\iota}^\top \\ \mathbb{E}A_{i\iota} & \mathbb{E}A_{i\iota}^2 & \mathbb{E}A_{i\iota}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}A_{i\iota}O_{i\iota}^\top \\ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}A_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}^2 & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}O_{i\iota}^\top \\ \mathbb{E}O_{i\iota} & \mathbb{E}A_{i\iota}O_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}O_{i\iota} & \mathbb{E}O_{i\iota}O_{i\iota}^\top \end{pmatrix}+O_p\left(N^{-1/2}\right)$$

We next calculate these expectations. By design,

$$\begin{aligned}
&\mathbb{E}A_{i\iota}=\mathbb{E}A_{i\iota}^2=p_\iota,\ \ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}=p,\ \ \mathbb{E}A_{i\iota}\overline{A}_{i\mathcal{N}_\iota}=(\mathbb{E}A_{i\iota})(\mathbb{E}\overline{A}_{i\mathcal{N}_\iota})=p^2,\\
&\mathbb{E}\overline{A}_{i\mathcal{N}_\iota}^2=\frac{1}{n_\iota^2}\mathbb{E}\left\{\sum_{j\in\mathcal{N}_\iota}A_{ij}\right\}^2=\frac{1}{n_\iota^2}\mathbb{E}\left\{\sum_{j\in\mathcal{N}_\iota}A_{ij}+\sum_{j_1\neq j_2;j_1,j_2\in\mathcal{N}_\iota}A_{ij_1}A_{ij_2}\right\}\\
&\qquad\quad=\frac{1}{n_\iota^2}[n_\iota p+n_\iota(n_\iota-1)p^2]=p^2+\frac{p(1-p)}{n_\iota},\\
&\mathbb{E}A_{i\iota}O_{i\iota}=p\mathbb{E}O_{i\iota},\ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota k}O_{i\iota k}=p\mathbb{E}O_{i\iota}.
\end{aligned}$$

Notice that the OLS estimator can be expressed as

$$\begin{pmatrix}\widehat{\alpha}_\iota\\ \widehat{\gamma}_\iota\\ \widehat{\theta}_\iota\\ \widehat{\beta}_\iota\end{pmatrix}=\begin{pmatrix}\alpha_\iota\\ \gamma_\iota\\ \theta_\iota\\ \beta_\iota\end{pmatrix}+\left(M_\iota^{ir\top}M_\iota^i\right)^{-1}M_\iota^{ir\top}e_\iota,$$

where $e_\iota=(e_{1\iota},\ldots,e_{N\iota})^\top\in\mathbb{R}^{N\times 1}$. Since $\widehat{\tau}$ depends on $\widehat{\gamma}_\iota$ and $\widehat{\theta}_\iota$, we need to calculate the second and third rows of $[N^{-1}(M_\iota^i)^\top M_\iota^i]^{-1}$. Using similar techniques in calculating the inverse of $(M_\iota^g)^\top M_\iota^g$, we can show that these two rows are equal to

$$\begin{pmatrix} -\dfrac{1}{1-p} & \dfrac{1}{p(1-p)} & 0 & 0\\ -\dfrac{n_\iota}{1-p} & 0 & \dfrac{n_\iota}{p(1-p)} & 0\end{pmatrix}.$$

It follows that

$$\widehat{\gamma}_\iota^i+\widehat{\theta}_\iota^i=\gamma_\iota+\theta_\iota+\frac{1+o_p(1)}{N}\sum_{i=1}^N\left\{\frac{A_{i\iota}-p}{(1-p)\,p}+\frac{n_\iota(\overline{A}_{i\iota}-p)}{p(1-p)}\right\}e_{i\iota}$$

and hence the MSE of the ATE estimator $\widehat{\tau}^i$ is given by

$$\begin{aligned}
&\mathrm{MSE}(\widehat{\tau}^i)\\
\asymp&\frac{1}{N}\mathrm{Var}\left[\sum_{\iota=1}^R\left\{\frac{A_{i\iota}-p}{(1-p)\,p}+\frac{n_\iota(\overline{A}_{i\iota}-p)}{p(1-p)}\right\}e_{i\iota}\right]\\
=&\frac{1}{N}\mathbb{E}\left[\sum_{\iota,\iota'=1}^R\left\{\frac{A_{i\iota}-p}{(1-p)p}+\frac{n_\iota(\overline{A}_{i\iota}-p)}{p(1-p)}\right\}\left\{\frac{A_{i\iota'}-p}{(1-p)p}+\frac{n_\iota'(\overline{A}_{i\iota'}-p)}{p(1-p)}\right\}e_{i\iota}e_{i\iota'}\right]\\
=&\frac{1}{N}\sum_{\iota=1}^R\frac{1}{(1-p)p}\mathbb{V}_{\iota\iota}+\frac{2}{N}\sum_{\iota=1}^R\sum_{\iota'\in\mathcal{N}_\iota}\frac{1}{p(1-p)}\mathbb{V}_{\iota\iota'}+\frac{1}{N}\sum_{\iota,\iota'}\sum_{j\in\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}}\frac{1}{p(1-p)}\mathbb{V}_{\iota\iota'}.
\end{aligned}$$

This yields the asymptotic variance of $\widehat{\tau}^i$.

Next, we consider the cluster-randomized design. For any region $\iota \in \mathcal{C}_j^0$, this region and all its neighboring regions receive the identical treatment. Similar to the global design, the estimator for $\gamma_\iota + \theta_\iota$ equals

$$\gamma_\iota + \theta_\iota + \frac{1+o_p(1)}{N} \sum_{i=1}^N \frac{A_{i\iota} - p}{p(1-p)} e_{i\iota}. \tag{3}$$

When $\iota \in \partial \mathcal{C}_j$, similar to the individual design, we can show that the inverse of $N^{-1}(M_\iota^r)^\top M_\iota^r$ is asymptotically equivalent to

$$\begin{aligned}
&\begin{pmatrix} 1 & \mathbb{E}A_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}O_{i\iota}^\top \\ \mathbb{E}A_{i\iota} & \mathbb{E}A_{i\iota}^2 & \mathbb{E}A_{i\iota}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}A_{i\iota}O_{i\iota}^\top \\ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}A_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}^2 & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}O_{i\iota}^\top \\ \mathbb{E}O_{i\iota} & \mathbb{E}A_{i\iota}O_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota}O_{i\iota} & \mathbb{E}O_{i\iota}O_{i\iota}^\top \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & 0\,0\,0 \\ \mathbb{E}A_{i\iota} & 1\,0\,0 \\ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & 0\,1\,0 \\ \mathbb{E}O_{i\iota} & 0\,0\,I \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0^\top \\ 0 & \mathbb{E}A_{i\iota} - (\mathbb{E}A_{i\iota})^2 & \mathbb{E}A_{i\iota}\overline{A}_{i\mathcal{N}_\iota} - \mathbb{E}A_{i\iota}\mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & 0^\top \\ 0 & \mathbb{E}A_{i\iota}\overline{A}_{i\mathcal{N}_\iota} - \mathbb{E}A_{i\iota}\mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}\overline{A}_{i\iota} - (\mathbb{E}\overline{A}_{i\iota})^2 & 0 \\ 0 & 0 & 0 & \mathrm{cov}(O_{i\iota}) \end{pmatrix}^{-1} \\
&\qquad \begin{pmatrix} 1 & \mathbb{E}A_{i\iota} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota} & \mathbb{E}O_{i\iota}^\top \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & -p & -p & -(\mathbb{E}O_{i\iota})^\top \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0^\top \\ 0 & p(1-p) & \frac{n_\iota^{(j)}}{n_\iota} p(1-p) & 0^\top \\ 0 & \frac{n_\iota^{(j)}}{n_\iota} p(1-p) & \mathrm{cov}(\bar{A}_{i\mathcal{N}_\iota}) & 0 \\ 0 & 0 & 0 & \mathrm{cov}(O_{i\iota}) \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0\,0\,0 \\ -p & 1\,0\,0 \\ p & 0\,1\,0 \\ -(\mathbb{E}O_{i\iota})^\top & 0\,0\,I \end{pmatrix}
\end{aligned}$$

where for any $1 \le k \le m$, $n_\iota^{(k)}$ denotes the number of neighbouring regions that belong to the kth cluster.

With some calculations, we can show that the second matrix on the second line equals

$$\begin{pmatrix} 1 & 0 & 0 & 0^\top \\ 0 & \frac{\sum_k (n_\ell^{(k)})^2}{p(1-p)\sum_{k\neq j}(n_\ell^{(k)})^2} & -\frac{n_\iota n_\iota^{(j)}}{p(1-p)\sum_{k\neq j} n_\iota^{(k)})^2} & 0^\top \\ 0 & -\frac{n_\iota n_\iota^{(j)}}{p(1-p)\sum_{k\neq j}(n_\iota^{(k)})^2} & \frac{n_\iota^2}{p(1-p)\sum_{k\neq j}(n_\iota^{(k)})^2} & 0^\top \\ 0 & 0 & 0 & \mathrm{cov}^{-1}(O_{i\iota}) \end{pmatrix}.$$

As such, the estimator for $\gamma_\iota + \theta_\iota$ equals

$$\begin{aligned}
\gamma_\iota + \theta_\iota + \frac{1+o_p(1)}{N} \sum_{i=1}^N \frac{\sum_{k\neq j}(n_\ell^{(k)})^2 - n_\iota^{(j)}\sum_{k\neq j} n_\iota^{(k)}}{p(1-p)\sum_{k\neq j}(n_\ell^{(k)})^2} (A_i^{(j)} - p) e_{i\iota} \\
+ \frac{1+o_p(1)}{N} \sum_{i=1}^N \frac{n_\iota \sum_{k\neq j} n_\iota^{(k)}}{\sum_{k\neq j}(n_\ell^{(k)})^2 p(1-p)} (\bar{A}_{i\iota} - p) e_{i\iota} \\
= \gamma_\iota + \theta_\iota + \frac{1+o_p(1)}{N} \sum_{i=1}^N \frac{A_i^{(j)} - p}{p(1-p)} e_{i\iota}
\end{aligned}$$

$$+\frac{1+o_p(1)}{N}\sum_{i=1}^{N}\frac{\sum_{k_1\neq j} n_\iota^{(k_1)}\sum_{k_2\neq j} n_\iota^{(k_2)}}{p(1-p)\sum_{k\neq j}(n_\ell^{(k)})^2}(A_i^{(k_2)}-p)e_{i\iota}.$$

Consequently,

$$\widehat{\tau}^c=\tau^c+\frac{1+o_p(1)}{N}\sum_{i=1}^{N}\sum_{j=1}^{m}\frac{A_i^{(j)}-p}{p(1-p)}\sum_{\iota\in\mathcal{C}_j} e_{i\iota}$$
$$+\frac{1+o_p(1)}{N}\sum_{i=1}^{N}\sum_{j=1}^{m}\sum_{\iota\in\partial\mathcal{C}_j}\frac{\sum_{k_1\neq j} n_\iota^{(k_1)}\sum_{k_2\neq j} n_\iota^{(k_2)}}{\sum_{k\neq j}(n_\ell^{(k)})^2 p(1-p)}(A_i^{(k_2)}-p)e_{i\iota}.$$

Its MSE is asymptotically equivalent to

$$\frac{1}{N}\sum_{j=1}^{m}\frac{\sum_{\iota,\iota'\in\mathcal{C}_j}\mathbb{V}_{\iota\iota'}}{q^{(j)}}+\frac{2}{N}\sum_{j=1}^{m}\sum_{j'\neq j}\sum_{\iota\in\mathcal{C}_j}\sum_{\iota'\in\partial\mathcal{C}_{j'}}\frac{\sum_{k_1\neq j'} n_{\iota'}^{(k_1)} n_{\iota'}^{(j)}}{\sum_{k\neq j'}(n_{\iota'}^{(k)})^2 q^{(k)}}\mathbb{V}_{\iota\iota'}$$
$$+\frac{1}{N}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sum_{\iota\in\partial\mathcal{C}_{j_1}}\sum_{\iota'\in\partial\mathcal{C}_{j_2}}\frac{\sum_{k\neq j_1,j_2} n_\iota^{(k)} n_{\iota'}^{(k)} q^{(k)}\sum_{k\neq j_1} n_\iota^{(k)}\sum_{k\neq j_2} n_{\iota'}^{(k)}}{\sum_{k\neq j_1}(n_\iota^{(k)})^2 q^{(k)}\sum_{k\neq j_2}(n_{\iota'}^{(k)})^2 q^{(k)}}\mathbb{V}_{\iota\iota'}.$$

This completes the proof.

3.2. *Proof of Theorem 2.2.* By setting $q=q_j=q^{(j)}=0.5$, it follows from Theorem S.1 that

$$\mathrm{MSE}\left(\widehat{\tau}^g\right)\asymp\frac{4}{N}\sum_{\iota=1}^{R}\sum_{\iota'=1}^{R}\mathbb{V}_{\iota\iota'},$$
$$\mathrm{MSE}\left(\widehat{\tau}^i\right)\asymp\frac{4}{N}\left\{\sum_{\iota=1}^{R}\mathbb{V}_{\iota\iota}+2\sum_{\iota=1}^{R}\sum_{\iota'\in\mathcal{N}_\iota}\mathbb{V}_{\iota\iota'}+\sum_{\iota,\iota'=1}^{R}|\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}|\,\mathbb{V}_{\iota\iota'}\right\},$$
$$\mathrm{MSE}(\widehat{\tau}^c)\asymp\frac{4}{N}\sum_{j=1}^{m}\left\{\sum_{\iota,\iota'\in\mathcal{C}_j}\mathbb{V}_{\iota\iota'}+2\sum_{\iota\in\mathcal{C}_j}\sum_{\iota'\in\mathcal{N}_{\mathcal{C}_j}}\omega_{\iota'}n_{\iota'}^{(j)}\mathbb{V}_{\iota\iota'}+\sum_{\iota,\iota'\in\mathcal{N}_{\mathcal{C}_j}}\omega_\iota\omega_{\iota'}n_\iota^{(j)}n_{\iota'}^{(j)}\mathbb{V}_{\iota\iota'}\right\},$$

where $\omega_\iota=\sum_{k\neq j} n_\iota^{(k)}/\sum_{k\neq j}(n_\iota^{(k)})^2$ for any ι.

According to the Cauchy-Schwarz inequality, we obtain that $\mathbb{V}_{\iota\iota'}\leq 0.5\mathbb{V}_{\iota\iota}+0.5\mathbb{V}_{\iota'\iota'}$ for any ι and ι'. It follows that

$$\frac{\mathrm{MSE}(\widehat{\tau}^i)}{\mathrm{MSE}(\widehat{\tau}^g)}\lesssim\frac{\sum_{\iota=1}^{R}\mathbb{V}_{\iota\iota}}{\sum_{\iota,\iota'=1}^{R}\mathbb{V}_{\iota\iota'}}+\frac{\sum_{\iota=1}^{R}\sum_{\iota'\in\mathcal{N}_\iota}(\mathbb{V}_{\iota\iota}+\mathbb{V}_{\iota'\iota'})}{\sum_{\iota,\iota'=1}^{R}\mathbb{V}_{\iota\iota'}}+\frac{0.5\sum_{\iota,\iota'}|\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}|(\mathbb{V}_{\iota\iota}+\mathbb{V}_{\iota'\iota'})}{\sum_{\iota,\iota'=1}^{R}\mathbb{V}_{\iota\iota'}}.$$

By definition, the first term on the right-hand-side (RHS) equals $1/\nu$. By symmetry, we have $\sum_{\iota=1}^{R}\sum_{\iota'\in\mathcal{N}_\iota}\mathbb{V}_{\iota'\iota'}=\sum_{\iota'=1}^{R}\sum_{\iota\in\mathcal{N}_{\iota'}}\mathbb{V}_{\iota'\iota'}$. Since $|\mathcal{N}_\iota|=n_\iota\leq r$, the second term on the RHS is upper bounded by $2r/\nu$. Finally, consider the third term on the RHS. Similarly, by leveraging symmetry, the numerator is upper bounded by $\sum_{\iota,\iota'}|\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}|\mathbb{V}_{\iota\iota}$. For a specific ι, consider the sum $\sum_{\iota'}|\mathcal{N}_\iota\cap\mathcal{N}_{\iota'}|$. It essentially represents the cumulative count of neighbors across regions within $\mathcal{N}_\iota$. Hence, it can be upper bounded by r^2/ν. Combining the three pieces together yields the desired upper bound for the ratio $\mathrm{MSE}(\widehat{\tau}^i)/\mathrm{MSE}(\widehat{\tau}^g)$.

Next, consider the ratio $\mathrm{MSE}(\widehat{\tau}^c)/\mathrm{MSE}(\widehat{\tau}^g)$. When $\omega = O(1)$, we have $\omega_\iota n_\iota^{(j)} = O(1)$ for any ι and j. It follows that

$$N\mathrm{MSE}(\widehat{\tau}^c) = O\Big[\sum_{j=1}^m \Big(\sum_{\iota,\iota'\in\mathcal{C}_j} \mathbb{V}_{\iota\iota'} + 2\sum_{\iota\in\mathcal{C}_j}\sum_{\iota'\in\mathcal{N}_{\mathcal{C}_j}} n_{\iota'}^{(j)}\mathbb{V}_{\iota\iota'} + \sum_{\iota,\iota'\in\mathcal{N}_{\mathcal{C}_j}} n_\iota^{(j)} n_{\iota'}^{(j)}\mathbb{V}_{\iota\iota'}\Big)\Big]$$
$$= O\Big(\sum_j \sum_{\iota,\iota'\in\mathcal{C}_j\cup\mathcal{N}_{\mathcal{C}_j}} \mathbb{V}_{\iota\iota'}\Big).$$

The proof is hence completed.

3.3. *Proof of Theorem S.2*. Again, we first consider the global design. Notice that the propensity score is correctly specified by design. According to the doubly robustness property, the estimator is unbiased to the ATE. Consequently, its MSE equals its variance. Notice that the ATE estimator can be represented as

$$\frac{1}{N}\sum_{k=1}^K\sum_{\iota=1}^R\sum_{t\in\mathcal{I}_k}\nu_{DR}(a,\iota,t,\widehat{h}_\iota^{(k)},\pi_\iota^g) = \frac{1}{N}\sum_{k=1}^K\sum_{\iota=1}^R\sum_{t\in\mathcal{I}_k}\Big\{\sum_{a=0}^1(-1)^{a+1}\frac{\mathbb{I}(A_i=a)}{pa+(1-p)(1-a)}e_{i,\iota}$$
$$(4)\quad \sum_{a=0}^1(-1)^{a+1}\frac{\mathbb{I}(A_i=a)-pa-(1-p)(1-a)}{pa+(1-p)(1-a)}[\widehat{h}_\iota^{(k)}(a,a,O_{i,\iota},\overline{O}_{i,\iota}) - h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})]$$
$$+\sum_{a=0}^1(-1)^{a+1}h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})\Big\}.$$

Notice that the first term on the RHS of the first line, the second line as well as the last line are mutually uncorrelated. As such, the MSE of the ATE estimator equals

$$\mathbb{E}\Big\{\frac{1}{N}\sum_{k=1}^K\sum_{\iota=1}^R\sum_{t\in\mathcal{I}_k}\sum_{a=0}^1(-1)^{a+1}\frac{\mathbb{I}(A_i=a)-pa-(1-p)(1-a)}{pa+(1-p)(1-a)}[\widehat{h}_\iota^{(k)}(a,a,O_{i,\iota},\overline{O}_{i,\iota}) - h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})]\Big\}^2$$
$$+\mathbb{E}\Big\{\frac{1}{N}\sum_{i=1}^N\sum_{\iota=1}^R\frac{\mathbb{I}(A_t=1)-p}{p(1-p)}e_{i,\iota}\Big\}^2 + \frac{\sigma_O^2}{N}.$$

With some calculations, the first term on the second line equals

$$\frac{1}{Np(1-p)}\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota,\iota'}.$$

It remains to show that the first line can be upper bounded by $[Np(1-p)]^{-1}RO(\sum_{\iota=1}^R\delta_{N,\iota}^2)$. By Cauchy-Schwarz inequality, the first line is upper bounded by

$$(5)\quad \begin{aligned} &\frac{2K}{N^2}\sum_{k=1}^K\mathbb{E}\Big\{\sum_{\iota=1}^R\sum_{t\in\mathcal{I}_k}\frac{A_t-p}{p}[\widehat{h}_\iota^{(k)}(1,1,O_{i,\iota},\overline{O}_{i,\iota}) - h_\iota(1,1,O_{i,\iota},\overline{O}_{i,\iota})]\Big\}^2 \\ &+\frac{2K}{N^2}\sum_{k=1}^K\mathbb{E}\Big\{\sum_{\iota=1}^R\sum_{t\in\mathcal{I}_k}\frac{A_t-p}{1-p}[\widehat{h}_\iota^{(k)}(0,0,O_{i,\iota},\overline{O}_{i,\iota}) - h_\iota(0,0,O_{i,\iota},\overline{O}_{i,\iota})]\Big\}^2 \end{aligned}$$

Notice that for each k and any $t_1\neq t_2\in\mathcal{I}_k$, $a\in\{0,1\}$, the variables

$$\tfrac{A_{i_1}-p}{p}[\widehat{h}_\iota^{(k)}(a,a,O_{i_1,\iota},\overline{O}_{t_1,\iota}) - h_\iota(a,a,O_{i_1,\iota},\overline{O}_{t_1,\iota})]$$
$$\text{and } \tfrac{A_{i_2}-p}{p}[\widehat{h}_\iota^{(k)}(a,a,O_{i_2,\iota},\overline{O}_{t_2,\iota}) - h_\iota(a,a,O_{i_2,\iota},\overline{O}_{t_2,\iota})]$$

are uncorrelated. With some calculations, it can be shown that (5) is upper bounded by $[Np(1-p)]^{-1}RO(\sum_{\iota=1}^{R}\delta_{N,\iota}^2)$. The proof for the global design is hence completed.

We next consider the individual-randomized design. Similar to (4), the ATE estimator can be rewritten as

$$\frac{1}{N}\sum_{\iota=1}^{R}\sum_{i=1}^{N}\sum_{a=0}^{1}(-1)^{a+1}h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})$$
$$+\frac{1}{N}\sum_{\iota=1}^{R}\sum_{i=1}^{N}\sum_{a=0}^{1}(-1)^{a+1}\frac{\mathbb{I}(A_{i,\iota}=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_{i,j}=a)}{p^{1+n_\iota}a+(1-p)^{1+n_\iota}(1-a)}e_{i,\iota}$$
$$+\frac{1}{N}\sum_{\iota=1}^{R}\sum_{k=1}^{K}\sum_{t\in\mathcal{I}_k}\sum_{a=0}^{1}(-1)^{a+1}\Big\{\frac{\mathbb{I}(A_{i,\iota}=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_{i,j}=a)}{p^{1+n_\iota}a+(1-p)^{1+n_\iota}(1-a)}-1\Big\}$$
$$\cdot[h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})-\widehat{h}_\iota^{(k)}(a,a,O_{i,\iota},\overline{O}_{i,\iota})],$$

and the three terms are mutually uncorrelated. Additionally, the variance of the first term equals $N^{-1}\sigma_O^2$. By the definitions of n_ι and $n_{\iota,\iota'}$, the variance of the second term can be shown to be equal to

$$\frac{1}{N}\sum_{\iota=1}^{R}\Big\{\frac{1}{p^{1+n_\iota}}+\frac{1}{(1-p)^{1+n_\iota}}\Big\}\mathbb{V}_{\iota,\iota}+\frac{1}{N}\sum_{\iota\neq\iota'}\Big\{\frac{1}{p^{n_{\iota,\iota'}}}+\frac{1}{(1-p)^{n_{\iota,\iota'}}}\Big\}\mathbb{V}_{\iota,\iota'}\mathbb{I}(n_{\iota\iota'}>0).$$

It remains to upper bound the variance of the third term. Notice that for any two regions ι_1,ι_2 such that $\mathcal{N}_{\iota_1}\cap\mathcal{N}_{\iota_2}=\emptyset$, the covariance between

$$Z_{\iota_1}^{(k)}=\sum_{a=0}^{1}(-1)^{a+1}\Big\{\frac{\mathbb{I}(A_{\iota_1}=a)\prod_{j\in\mathcal{N}_{\iota_1}}\mathbb{I}(A_j=a)}{p^{1+n_{\iota_1}}a+(1-p)^{1+n_{\iota_1}}(1-a)}-1\Big\}[h_{\iota_1}(a,a,O_{\iota_1},\overline{O}_{\iota_1})-\widehat{h}_{\iota_1}^{(k)}(a,a,O_{\iota_1},\overline{O}_{\iota_1})]$$

and

$$Z_{\iota_2}^{(k)}=\sum_{a=0}^{1}(-1)^{a+1}\Big\{\frac{\mathbb{I}(A_{\iota_2}=a)\prod_{j\in\mathcal{N}_{\iota_2}}\mathbb{I}(A_j=a)}{p^{1+n_{\iota_2}}a+(1-p)^{1+n_{\iota_2}}(1-a)}-1\Big\}[h_{\iota_2}(a,a,O_{\iota_2},\overline{O}_{\iota_2})-\widehat{h}_{\iota_2}^{(k)}(a,a,O_{\iota_2},\overline{O}_{\iota_2})]$$

equals zero. Additionally, for each region ι, the number of its neighbouring regions is bounded by r. Consequently, the number of its neighbour's neighbour is bounded by r^2. According to the Cauchy-Schwarz inequality, we have that

$$\mathrm{Var}(\sum_{\iota=1}^{R}Z_\iota^{(k)})=\sum_{\iota=1}^{R}\mathrm{Var}(Z_\iota^{(k)})+\sum_{\substack{\iota_1\neq\iota_2\\ \mathcal{N}_{\iota_1}\cap\mathcal{N}_{\iota_2}\neq\emptyset}}\mathrm{cov}(Z_{\iota_1}^{(k)},Z_{\iota_2}^{(k)})$$
$$\leq\sum_{\iota=1}^{R}\mathrm{Var}(Z_\iota^{(k)})+\sum_{\substack{\iota_1\neq\iota_2\\ \mathcal{N}_{\iota_1}\cap\mathcal{N}_{\iota_2}\neq\emptyset}}\frac{\mathrm{Var}(Z_{\iota_1}^{(k)})+\mathrm{Var}(Z_{\iota_2}^{(k)})}{2}$$
$$=O(r^2\sum_{\iota=1}^{R}\mathrm{Var}(Z_\iota^{(k)}))=O\Big(\frac{r^2}{p^{r+1}}\sum_{\iota=1}^{R}\delta_{N,\iota}^2+\frac{r^2}{(1-p)^{r+1}}\sum_{\iota=1}^{R}\delta_{N,\iota}^2\Big).$$

Combining this together with the analysis in (5) yields the desired upper bound for the variance of the third term. The proof for the individual-randomized design is hence completed.

Finally, consider the cluster-randomized design. Similarly, we can decompose the ATE estimator into the sum of three terms, given by

$$\frac{1}{N}\sum_{\iota=1}^{R}\sum_{i=1}^{N}\sum_{a=0}^{1}(-1)^{a+1}h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})$$
$$+\frac{1}{N}\sum_{\iota=1}^{R}\sum_{i=1}^{N}\sum_{a=0}^{1}(-1)^{a+1}\frac{\mathbb{I}(A_{i,\iota}=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_{i,j}=a)}{\pi_\iota^c(a)}e_{i,\iota}$$
$$+\frac{1}{N}\sum_{\iota=1}^{R}\sum_{k=1}^{K}\sum_{t\in\mathcal{I}_k}\sum_{a=0}^{1}(-1)^{a+1}\Big\{\frac{\mathbb{I}(A_{i,\iota}=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_{i,j}=a)}{\pi_\iota^c(a)}-1\Big\}$$
$$[h_\iota(a,a,O_{i,\iota},\overline{O}_{i,\iota})-\widehat{h}_\iota^{(k)}(a,a,O_{i,\iota},\overline{O}_{i,\iota})].$$

Again, the variance of the first term is given by $N^{-1}\sigma_O^2$. As for the second term, consider two regions ι and ι'. By definition, the covariance between

$$\sum_{a=0}^{1}(-1)^{a+1}\frac{\mathbb{I}(A_\iota=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_j=a)}{\pi_\iota^c(a)}e_\iota \text{ and } \sum_{a=0}^{1}(-1)^{a+1}\frac{\mathbb{I}(A_{\iota'}=a)\prod_{j\in\mathcal{N}_{\iota'}}\mathbb{I}(A_j=a)}{\pi_{\iota'}^c(a)}e_{\iota'}$$

depends on the value $m_{\iota\iota'}$. Specifically, when $m_{\iota\iota'}=0$, their covariance equals 0. Otherwise, their covariance equals $[p^{-m_{\iota\iota'}}+(1-p)^{-m_{\iota\iota'}}]\mathbb{V}_{\iota\iota'}$. Consequently, it remains to bound the last term.

Similar to (5), it suffices to bound the variance of

$$\sum_{\iota=1}^{R}Z_\iota^{(a,k)}=\sum_{\iota=1}^{R}\Big\{\frac{\mathbb{I}(A_\iota=a)\prod_{j\in\mathcal{N}_\iota}\mathbb{I}(A_j=a)}{\pi_\iota^c(a)}-1\Big\}[h_\iota(a,a,O_\iota,\overline{O}_\iota)-\widehat{h}_\iota^{(k)}(a,a,O_\iota,\overline{O}_\iota)]$$

for each a and k. Using Cauchy-Schwarz inequality again, its variance can be upper bounded by

$$2\text{Var}(\sum_{k=1}^{m}\sum_{\iota\in\mathcal{C}_k^0}Z_\iota^{(a,k)})+2\text{Var}(\sum_{k=1}^{m}\sum_{\iota\in\partial\mathcal{C}_k}Z_\iota^{(a,k)}).$$

For any two regions $\iota,\iota'\in\mathcal{C}_k^0$, both these regions and their neighbours receive the same treatment assignment. On the contrary, for any $\iota\in\mathcal{C}_k^0$ and $\iota'\in\mathcal{C}_{k'}^0$ such that $k\neq k'$, the two regions and their neighbours receive independent treatments. As such, the first term is given by

$$2\sum_{k=1}^{m}\text{Var}(\sum_{\iota\in\mathcal{C}_k^0}Z_\iota^{(a,k)}),$$

which can be upper bounded by $O(p^{-1}(1-p)^{-1}\sum_{k=1}^m|\mathcal{C}_k^0|\sum_{\iota\in\mathcal{C}_k^0}\delta_{N,\iota}^2)$, according to the Cauchy-Schwarz inequality.

As for the second term, using Cauchy-Schwarz inequality again, we can show that it is upper bounded by

$$2\sum_{k=1}^{m}\text{Var}(\sum_{\iota\in\partial\mathcal{C}_k}Z_\iota^{(a,k)})+2\sum_{\substack{k_1\neq k_2\\ \mathcal{N}_{\mathcal{C}_{k_1}}\cap\mathcal{N}_{\mathcal{C}_{k_2}}\neq\emptyset}}\text{cov}(\sum_{\iota\in\partial\mathcal{C}_{k_1}}Z_\iota^{(a,k_1)},\sum_{\iota\in\partial\mathcal{C}_{k_2}}Z_\iota^{(a,k_2)})$$
$$\leq 2\sum_{k=1}^{m}|\partial\mathcal{C}_k|\sum_{\iota\in\partial\mathcal{C}_k}\text{Var}(Z_\iota^{(a,k)})+2m'\sum_{k=1}^{m}|\partial\mathcal{C}_k|\sum_{\iota\in\partial\mathcal{C}_k}\text{Var}(Z_\iota^{(a,k)}).$$

This together with the definition of $\delta_{N,\iota}^2$ and r' yields the desired result. The proof is hence completed.

3.4. *Proof of Theorem 2.4.* According to Theorem S.2, to establish the first assertion, it suffices to establish an upper bound for

$$\frac{\sum_{\iota=1}^R 2^{n_\iota+2}\mathbb{V}_{\iota\iota}}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}+\frac{\sum_{\iota\neq\iota'} 2^{n_{\iota\iota'}+1}\mathbb{V}_{\iota\iota'}\mathbb{I}(n_{\iota\iota'}>0)}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}.$$

Under the condition that $n_\iota \le r$, the first term is upper bounded by $2^r/\nu$. As for the second term, by symmetry and Cauchy-Schwarz inequality, we obtain that

$$\frac{\sum_{\iota\neq\iota'} 2^{n_{\iota\iota'}+1}\mathbb{V}_{\iota\iota'}\mathbb{I}(n_{\iota\iota'}>0)}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}\le\frac{\sum_{\iota\neq\iota'} 2^{n_{\iota\iota'}}(\mathbb{V}_{\iota\iota}+\mathbb{V}_{\iota'\iota'})\mathbb{I}(n_{\iota\iota'}>0)}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}=\frac{\sum_{\iota\neq\iota'} 2^{n_{\iota\iota'}-1}\mathbb{V}_{\iota\iota}\mathbb{I}(n_{\iota\iota'}>0)}{\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}.$$

Consider a given ι. The set of regions $\{\iota' : n_{\iota\iota'} > 0\}$ include its neighbours $\mathcal{N}_\iota$ as well as its neighbours' neighbours. Consequently, the cardinality of this set is upper bounded by $(r+1)r$. Meanwhile, by definition, $n_{\iota\iota'}$ is upper bounded by $\min(n_\iota, n_{\iota'}) \le r+1$. This yields the following upper bound for the second term: $\nu^{-1}r(r+1)2^r$. Combining this with the upper bound of the first term proves the first assertion.

Next, consider the second assertion. Similarly, it suffices to establish an upper bound for

$$\frac{\sum_{\iota\iota'} 2^{m_{\iota\iota'}+1}\mathbb{V}_{\iota\iota'}\mathbb{I}(m_{\iota\iota'}>0)}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}},$$

or

$$\frac{\sum_{\iota\iota'} 2^{m_{\iota\iota'}+1}\mathbb{V}_{\iota\iota}\mathbb{I}(m_{\iota\iota'}>0)}{4\sum_{\iota,\iota'=1}^R \mathbb{V}_{\iota\iota'}}, \tag{6}$$

by Cauchy-Schwatrz inequality. Denote $\mathcal{N}_\iota^o = \mathcal{N}_\iota \cup \{\iota\}$.

Then we have

$$\begin{aligned}&\sum_{\iota,\iota'} 2^{m_{\iota\iota'}}\mathbb{V}_{\iota\iota}\mathbb{I}(m_{\iota\iota'}>0)\\ \le &\sum_{\iota\in\mathcal{R}/\mathcal{R}_1}\mathbb{V}_{\iota\iota}\cdot O(cr)+\sum_{\iota\in\mathcal{R}_1}\mathbb{V}_{\iota\iota}\cdot 2^{r_c}\cdot crr_c.\end{aligned}$$

This completes the proof.

3.5. *Proof of Proposition 2.* The IE is

$$\begin{aligned}&\sum_{\tau=1}^M\sum_{\iota=1}^R\beta_{\iota\tau}^\top\sum_{k=1}^{\tau-1}\left(\prod_{j=k+1}^{\tau-1}B_{\iota j}\right)\{\Gamma_{\iota k}+\Theta_{\iota k}\}\\ =&\sum_{\tau=1}^M\sum_{\iota=1}^R\sum_{k=1}^{\tau-1}\beta_{\iota\tau}^\top\left(\prod_{j=k+1}^{\tau-1}B_{\iota j}\right)\{\Gamma_{\iota k}+\Theta_{\iota k}\}\\ =&\sum_{k=1}^M\sum_{\iota=1}^R\left\{\sum_{\tau=k+1}^M\beta_{\iota\tau}^\top\left(\prod_{j=k+1}^{\tau-1}B_{\iota j}\right)\right\}\Big\{\Gamma_{\iota k}+\Theta_{\iota k}\Big\}.\end{aligned}$$

3.6. *Proofs of Theorem S.3*. For conciseness, we denote $\widehat{\tau} = \widehat{\tau}_{OLS}$. Denote $b_{\iota jt}^{\top} = \sum_{k=t+1}^{M} \beta_{\iota k}^{\top} \prod_{\substack{j_1=t+1 \\ j_1 \neq j}}^{k-1} B_{\iota j_1}$, $\mathbb{B}_{\iota jt} = \prod_{k=j+1}^{t-1} B_{\iota k}$ and $\prod_{j=k}^{t} B_{\iota j} = 1$ if $t < k$, and $C_{\iota t} = \sum_{j=1}^{t-1} \mathbb{B}_{\iota jt} \Gamma_{\iota j}^{g} a_{jt}$, which are parameters to describe accumulated effects due to the indirect effects. We also introduce the following noise quantity,

$$v_{i\iota jt} = b_{\iota jt}^{\top} E_{i\iota j} \Gamma_{\iota t}^{g} + \mathbb{B}_{\iota jt} \Gamma_{\iota t}^{g} e_{i\iota j}.$$

1. $MSE(\widehat{\tau}^{g})$.

Denote

$$M_{\iota t}^{g} = \begin{pmatrix} 1 & A_{1t} & O_{1\iota t}^{\top} \\ 1 & A_{2t} & O_{2\iota t}^{\top} \\ \vdots & \vdots & \vdots \\ 1 & A_{Nt} & O_{N\iota t}^{\top} \end{pmatrix}.$$

The OLS coefficient estimates are

$$\begin{pmatrix} \widehat{\alpha}_{\iota t} \\ \widehat{\gamma}_{\iota t}^{g} \\ \widehat{\beta}_{\iota t} \end{pmatrix} = \begin{pmatrix} \alpha_{\iota t} \\ \gamma_{\iota t}^{g} \\ \beta_{\iota t} \end{pmatrix} + \frac{1}{N} \left((M_{\iota t}^{g})^{\top} M_{\iota t}^{g} \right)^{-1} \cdot \frac{1}{N} (M_{\iota t}^{g})^{\top} e_{\iota t},$$

$$\left(\widehat{\Lambda}_{\iota t} \; \widehat{\Gamma}_{\iota t}^{g} \; \widehat{B}_{\iota t} \right) = \left(\Lambda_{\iota t} \; \Gamma_{\iota t}^{g} \; B_{\iota t} \right) + \frac{1}{N} E_{\iota t}^{\top} M_{\iota t}^{g} \cdot \frac{1}{N} \left((M_{\iota t}^{g})^{\top} M_{\iota t}^{g} \right)^{-1},$$

where $e_{\iota t} = (e_{1\iota t}, \ldots, e_{N\iota t})^{\top} \in \mathbb{R}^{N\times 1}$ and $E_{\iota t} = (E_{1\iota t}, \ldots, E_{N\iota t})^{\top} \in \mathbb{R}^{N\times d}$. We remark that as $\widehat{\tau}^{g}$ is derived based on the OLS estimates, and the key problem is to calculate the inverse of the Gram matrix.

From (14), we can derive that for $i = 1, \ldots, N$, $\iota = 1, \cdots, R$, $t = 1, \ldots, M$,

$$O_{i\iota t} = \Lambda_{\iota,t-1}^{*} + \mathbb{B}_{\iota 0t} O_{i\iota 1} + \sum_{j=1}^{t-1} \mathbb{B}_{\iota jt} \Gamma_{\iota j}^{g} A_{ij} + E_{i,\iota,t-1}^{*} \tag{7}$$

where $\mathbb{B}_{\iota jt} = \prod_{k=j+1}^{t-1} B_{\iota k}$, $\Lambda_{\iota,t-1}^{*} = \sum_{j=1}^{t} \mathbb{B}_{\iota jt} \Lambda_{\iota j}$ and $E_{i,\iota,t-1}^{*} = \sum_{j=1}^{t-1} \mathbb{B}_{\iota jt} E_{\iota j}$. By Large Number Theorem,

$$\frac{1}{N} (M_{\iota t}^{g})^{\top} M_{\iota t}^{g} = \begin{pmatrix} 1 & \mathbb{E} A_{it} & \mathbb{E} O_{i\iota t}^{\top} \\ \mathbb{E} A_{it} & \mathbb{E} A_{it}^{2} & \mathbb{E} A_{it} O_{i\iota t}^{\top} \\ \mathbb{E} O_{i\iota t} & \mathbb{E} A_{it} O_{i\iota t} & \mathbb{E} O_{i\iota t} O_{i\iota t}^{\top} \end{pmatrix} + O_p\left(N^{-1/2}\right).$$

Recall that given t, $A_{it} \overset{\text{i.i.d}}{\sim} \text{Bernoulli}(p)$, $\text{corr}(A_{it}, A_{it'}) = a_{tt'}$ which are independent of $\{(e_{i\iota t}, E_{i\iota t}^{\top})\}_{i,\iota,t}$. Denote

$$O_{i\iota t}^{0} = \mathbb{B}_{\iota 0t} (O_{i\iota 1} - \mathbb{E} O_{i\iota 1}) + E_{\iota t-1}^{*}, \quad S_{\iota t} = \text{Var}(O_{i\iota t}^{0})^{-1}.$$

Then direct algebra gives

$$\left\{ \frac{1}{N} (M_{\iota t}^{g})^{\top} M_{\iota t}^{g} \right\}^{-1} = \begin{pmatrix} & * & * & * \\ -\frac{1}{1-p} - C_{\iota t}^{\top} S_{\iota t} \mathbb{E} O_{i\iota t} & \frac{1}{p(1-p)} + C_{\iota t}^{\top} S_{\iota t} C_{\iota t} & -C_{\iota t}^{\top} S_{\iota t} \\ -S_{\iota t} \mathbb{E} O_{i\iota t} & -S_{\iota t} C_{\iota t} & S_{\iota t} \end{pmatrix} + O_p\left(N^{-1/2}\right).$$

Then

$$\left(\widehat{\gamma}_{\iota t}^{g} + \widehat{C}_{\iota t}^{g,\top} \widehat{\Gamma}_{\iota t}^{g} \right) - \left(\gamma_{\iota t}^{g} + C_{\iota t}^{\top} \Gamma_{\iota t}^{g} \right)$$

$$\asymp(\widehat{\gamma}_{\iota t}^{g}-\gamma_{\iota t}^{g})+C_{\iota t}^{\top}\left(\widehat{\Gamma}_{\iota t}^{g}-\Gamma_{\iota t}^{g}\right)+\sum_{j=t+1}^{M-1}b_{\iota jt}^{\top}\left(\widehat{B}_{\iota j}^{g}-B_{\iota j}\right)\Gamma_{\iota t}^{g}+\sum_{\tau=t+1}^{M}\left(\widehat{\beta}_{\iota\tau}^{g}-\beta_{\iota\tau}\right)^{\top}\mathbb{B}_{\iota t\tau}\Gamma_{\iota t}^{g}$$

$$\asymp\left\{\left(\frac{1}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{\iota t}\right)(A_{it}-p)-C_{\iota t}^{\top}S_{\iota t}(O_{i\iota t}-\mathbb{E}O_{i\iota t})\right\}u_{i\iota t}$$

(8)

$$+\sum_{k=t+1}^{M}\left\{S_{\iota k}(O_{i\iota k}-\mathbb{E}O_{i\iota k})-S_{\iota t}C_{\iota t}(A_{ik}-p)\right\}^{\top}v_{i\iota k},$$

where $b_{\iota jk}^{\top}=\sum_{\tau=j+1}^{M}\beta_{\iota\tau}^{\top}\prod_{\substack{j_1=k+1\\ j_1\neq j}}^{\tau-1}B_{\iota j_1}$. Denote

$$\sigma_{OLS}^2=\text{Var}\left[\sum_{t=1}^{M}\sum_{\iota=1}^{R}(O_{i\iota t}^{0})^{\top}\left\{-S_{\iota t}C_{\iota t}u_{i\iota t}+\sum_{k=t+1}^{M}S_{\iota k}v_{i\iota kt}\right\}\right].$$

Then we have

$$\begin{aligned}&\text{MSE}\left(\widehat{\tau}^{g}\right)-\sigma_{OLS}^{2}\\ \asymp&\mathbb{E}\left[\sum_{t=1}^{M}\sum_{\iota=1}^{R}\left\{\left(\frac{1}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{\iota t}\right)(A_{it}-p)-C_{\iota t}^{\top}S_{\iota t}\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Gamma_{\iota j}^{g}(A_{ij}-p)\right\}u_{i\iota t}\right.\\ &\left.+\sum_{t=1}^{M}\sum_{k=t+1}^{M}\sum_{\iota=1}^{R}\left\{S_{\iota k}\sum_{j=1}^{k-1}\mathbb{B}_{\iota jk}\Gamma_{\iota j}^{g}(A_{ij}-p)-S_{\iota k}C_{\iota k}(A_{ik}-p)\right\}^{\top}v_{i\iota k}\right]^{2}.\end{aligned}$$

Note that $u_{i\iota t}$ depends on $e_{i\iota t},E_{i\iota t}$ and $v_{i\iota kt}$ depends on $e_{i\iota k},E_{i\iota k}$. By the temporal independence of noises, we known that for $t\neq t',k\neq k'$ and $k>t$, $\text{cov}(u_{i\iota t},u_{i\iota t'})=0$, $\text{cov}(v_{i\iota k},v_{i\iota k'})=0$ and $\text{cov}(u_{i\iota t},v_{i\iota k})=0$. Hence

$$\text{MSE}\left(\widehat{\tau}^{g}\right)-\sigma_{OLS}^{2}\asymp\sum_{t=1}^{M}\mathbb{E}\left(\sum_{\iota=1}^{R}M_{1,\iota t}^{g}u_{i\iota t}\right)^{2}+\sum_{t=1}^{M}\sum_{k=t+1}^{M}\mathbb{E}\left\{\sum_{\iota=1}^{R}(M_{2,\iota k}^{g})^{\top}v_{i\iota k}\right\}^{2},$$

where

$$\begin{aligned}M_{1,\iota t}^{g}&=\left(\tfrac{1}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{\iota t}\right)(A_{it}-p)-C_{\iota t}^{\top}S_{\iota t}\textstyle\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Gamma_{\iota j}^{g}(A_{ij}-p),\\ M_{2,\iota k}^{g}&=S_{\iota k}\textstyle\sum_{j=1}^{k-1}\mathbb{B}_{\iota jk}\Gamma_{\iota j}^{g}(A_{ij}-p)-S_{\iota k}C_{\iota k}(A_{ik}-p).\end{aligned}$$

2. $MSE(\widehat{\tau}^{i})$.

Similarly, in model (13), we can derive that for $i=1,\dots,N$, $\iota=1,\cdots,R$, $t=1,\dots,M$,

$$O_{i\iota t}=\Lambda_{\iota,t-1}^{*}+\mathbb{B}_{0\iota t}O_{i\iota 1}+\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}(\Gamma_{\iota j}A_{i\iota j}+\Theta_{\iota j}\overline{A}_{i\mathcal{N}_{\iota}j})+E_{i,\iota,t-1}^{*}$$

where $\Lambda_{\iota,t-1}^{*}=\sum_{j=1}^{t}\mathbb{B}_{\iota jt}\Lambda_{\iota j}$ and $E_{i,\iota,t-1}^{*}=\sum_{j=1}^{t}\mathbb{B}_{\iota jt}E_{\iota j}$. Denote

$$M_{\iota t}=\begin{pmatrix}1 & A_{1\iota t} & \overline{A}_{1\mathcal{N}_{\iota}t} & O_{1\iota t}^{\top}\\ 1 & A_{2\iota t} & \overline{A}_{2\mathcal{N}_{\iota}t} & O_{2\iota t}^{\top}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & A_{N\iota t} & \overline{A}_{N\mathcal{N}_{\iota}t} & O_{N\iota t}^{\top}\end{pmatrix}.$$

Then we have

$$\frac{1}{N}M_{\iota t}^{\top}M_{\iota t}=\begin{pmatrix}1 & \mathbb{E}A_{i\iota t} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t} & \mathbb{E}O_{i\iota t}^{\top}\\ \mathbb{E}A_{i\iota t} & \mathbb{E}A_{i\iota t}^2 & \mathbb{E}A_{i\iota t}\overline{A}_{i\mathcal{N}_\iota t} & \mathbb{E}A_{i\iota t}O_{i\iota t}^{\top}\\ \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t}A_{i\iota t} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t}^2 & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t}O_{i\iota t}^{\top}\\ \mathbb{E}O_{i\iota t} & \mathbb{E}A_{i\iota t}O_{i\iota t} & \mathbb{E}\overline{A}_{i\mathcal{N}_\iota t}O_{i\iota t} & \mathbb{E}O_{i\iota t}O_{i\iota t}^{\top}\end{pmatrix}+O_p\left(N^{-1/2}\right)$$

In the following, when $A_{i\iota t}$ follows the individual-randomized design, we denote the above matrix as $N^{-1}(M_{\iota t}^{i})^{\top}M_{\iota t}^{i}$; when $A_{i\iota t}$ follows the cluster-randomized design, we denote the above matrix as $N^{-1}(M_{\iota t}^{c})^{\top}M_{\iota t}^{c}$. Further denote

$$C_{1,\iota,t}=\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Gamma_{\iota j}a_{jt},\quad C_{1,\iota,t}=\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Theta_{\iota j}a_{jt}.$$

Note that $C_{\iota t}=C_{1,\iota t}+C_{2,\iota t}$. Then it holds that

$$\begin{aligned}&\left\{\frac{1}{N}(M_{\iota t}^{i})^{\top}M_{\iota t}^{i}\right\}^{-1}\\ =&\begin{pmatrix} * & * & * & *\\ -\frac{1}{1-p}+C_{1,\iota,t}^{\top}S_{\iota t}\mathbb{E}O_{i\iota t}^{0} & \frac{1}{p(1-p)}+C_{2,\iota t}^{\top}S_{\iota t}C_{2,i\iota t} & C_{2,\iota t}^{\top}S_{\iota t}C_{1\iota t} & -C_{1,\iota,t}^{\top}S_{\iota t}\\ -\frac{n_\iota}{1-p}+C_{2,\iota,t}^{\top}S_{\iota t}\mathbb{E}O_{i\iota t}^{0} & C_{2,\iota t}^{\top}S_{\iota t}C_{1\iota t} & \frac{n_\iota}{p(1-p)}+C_{2,\iota t}^{\top}S_{\iota t}C_{2\iota t} & -C_{2,\iota,t}^{\top}S_{\iota t}\\ -S_{\iota t}\mathbb{E}O_{i\iota t}^{0} & -S_{\iota t}C_{1,\iota,t} & -S_{\iota t}C_{2,\iota,t} & S_{\iota t}\end{pmatrix}\\ &+O_p\left(N^{-1/2}\right).\end{aligned}$$

With these preliminaries, we have

$$\begin{aligned}&\left(\widehat{\gamma}_{\iota t}^{i}+\widehat{c}_{\iota t}^{i,\top}\widehat{\Gamma}_{\iota t}^{i}\right)-\left(\gamma_{\iota t}+c_{\iota t}^{\top}\Gamma_{\iota t}\right)\\ \asymp&\left(\widehat{\gamma}_{\iota t}^{i}-\gamma_{\iota t}+\widehat{\theta}_{\iota t}^{i}-\theta_{\iota t}\right)+C_{\iota t}^{\top}\left(\widehat{\Gamma}_{\iota t}^{i}-\Gamma_{\iota t}+\widehat{\Theta}_{\iota t}^{i}-\Theta_{\iota t}\right)\\ &+\sum_{j=t+1}^{M-1}b_{\iota jt}^{\top}\left(\widehat{B}_{\iota j}^{i}-B_{\iota j}\right)\Gamma_{\iota t}^{g}+\sum_{\tau=t+1}^{M}\left(\widehat{\beta}_{\iota\tau}^{i}-\beta_{\iota\tau}\right)^{\top}\mathbb{B}_{\iota t\tau}\Gamma_{\iota t}^{g}\\ \asymp&\left(\frac{1}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{1,\iota t}\right)(A_{i\iota t}-p)u_{i\iota t}\\ &+\left\{\left(\frac{n_\iota}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{2,\iota t}\right)(\overline{A}_{i\iota t}-p)-C_{\iota t}^{\top}S_{\iota t}(O_{i\iota t}-\mathbb{E}O_{i\iota t})\right\}u_{i\iota t}\\ &+\sum_{k=t+1}^{M}\left\{S_{\iota k}(O_{i\iota k}-\mathbb{E}O_{i\iota k})-S_{\iota k}C_{1,\iota t}(A_{i\iota k}-p)-S_{\iota k}C_{2,\iota t}(\overline{A}_{i\iota k}-p)\right\}^{\top}v_{i\iota kt}.\end{aligned}\tag{9}$$

Then we have

$$\mathrm{MSE}\left(\widehat{\tau}^{i}\right)-\sigma_{OLS}^{2}\asymp\sum_{t=1}^{M}\mathbb{E}\left(\sum_{\iota=1}^{R}M_{1,\iota t}^{i}u_{i\iota t}\right)^{2}+\sum_{t=1}^{M}\sum_{k=t+1}^{M}\mathbb{E}\left\{\sum_{\iota=1}^{R}(M_{2,\iota k}^{i})^{\top}v_{i\iota k}\right\}^{2},$$

where

$$\begin{aligned}M_{1,\iota t}^{i}=&\left(\tfrac{1}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{1,\iota t}\right)(A_{i\iota t}-p)-C_{\iota t}^{\top}S_{\iota t}\textstyle\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Gamma_{\iota j}(A_{i\iota j}-p)\\ &+\left(\tfrac{n_\iota}{p(1-p)}+C_{\iota t}^{\top}S_{\iota t}C_{2,\iota t}\right)(\overline{A}_{i\iota t}-p)-C_{\iota t}^{\top}S_{\iota t}\textstyle\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Theta_{\iota j}(\overline{A}_{i\iota j}-p),\end{aligned}$$

$$M^i_{2,\iota k} = S_{\iota k}\sum_{j=1}^{k-1}\mathbb{B}_{\iota jk}\Gamma_{\iota j}(A_{i\iota j}-p) - S_{\iota t}C_{1,\iota t}(A_{i\iota k}-p)$$
$$+S_{\iota k}\sum_{j=1}^{k-1}\mathbb{B}_{\iota jk}\Theta_{\iota j}(\overline{A}_{i\iota j}-p) - S_{\iota k}C_{2,\iota k}(\overline{A}_{i\iota k}-p).$$

3. $MSE(\widehat{\tau}^c)$.

For unit $\iota\in\mathcal{C}^0_{j_c}$, $\widehat{\gamma}^c_{\iota t}+\widehat{\theta}^c_{\iota t}+\widehat{c}^{c,\top}_{\iota t}(\widehat{\Gamma}^c_{\iota t}+\widehat{\Theta}^c_{\iota t})$ has the same expression as the global randomized design as in (8). We now derive the expression of $\widehat{\gamma}^c_{\iota t}+\widehat{\theta}^c_{\iota t}+\widehat{c}^{c,\top}_{\iota t}(\widehat{\Gamma}^c_{\iota t}+\widehat{\Theta}^c_{\iota t})$ for $\iota\in\partial\mathcal{C}_{j_c}$.

Denote

$$\eta_\iota = n_\iota^{-1}n_\iota^{(j_c)},\quad \zeta_\iota = n_\iota^{-2}\sum_{k_c}(n_\iota^{(k_c)})^2.$$

Then we compute the inverse of $N^{-1}(M^c_{\iota t})^\top M^c_{\iota t}$,

$$\left\{\frac{1}{N}(M^c_{\iota t})^\top M^c_{\iota t}\right\}^{-1}$$
$$=\begin{pmatrix} * & * & * & * \\ -\frac{\zeta_\iota-\eta_\iota}{\zeta_\iota-\eta_\iota^2}\frac{1}{1-p}+C^\top_{1,\iota,t}S_{\iota t}\mathbb{E}O^0_{i\iota t} & \frac{\zeta_\iota}{\zeta_\iota-\eta_\iota^2}\frac{1}{p(1-p)}+C^\top_{1,\iota t}S_{\iota t}C_{1,i\iota t} & -\frac{\eta_\iota}{\zeta_\iota-\eta_\iota^2}\frac{1}{p(1-p)}+C^\top_{2,\iota t}S_{\iota t}C_{1\iota t} & -C^\top_{1,\iota,t}S_{\iota t} \\ -\frac{1-\eta_\iota}{\zeta_\iota-\eta_\iota^2}\frac{1}{1-p}+C^\top_{2,\iota,t}S_{\iota t}\mathbb{E}O^0_{i\iota t} & -\frac{\eta_\iota}{\zeta_\iota-\eta_\iota^2}\frac{1}{p(1-p)}+C^\top_{2,\iota t}S_{\iota t}C_{1\iota t} & \frac{1}{\zeta_\iota-\eta_\iota^2}\frac{1}{p(1-p)}+C^\top_{2,\iota t}S_{\iota t}C_{2\iota t} & -C^\top_{2,\iota,t}S_{\iota t} \\ -S_{\iota t}\mathbb{E}O^0_{i\iota t} & -S_{\iota t}C_{1,\iota,t} & -S_{\iota t}C_{2,\iota,t} & S_{\iota t}\end{pmatrix}$$
$$+O_p\left(N^{-1/2}\right).$$

This gives

$$\left(\widehat{\gamma}^c_{\iota t}+\widehat{c}^{c,\top}_{\iota t}\widehat{\Gamma}^c_{\iota t}\right)-\left(\gamma_{\iota t}+c^\top_{\iota t}\Gamma_{\iota t}\right)$$
$$\asymp\left\{\left(\frac{\zeta_\iota-\eta_\iota}{\zeta_\iota^2-\eta_\iota}\frac{1}{p(1-p)}+C^\top_{\iota t}S_{\iota t}C_{1,\iota t}\right)(A_{i\iota t}-p)\right.$$
$$\left.+\left(\frac{1-\eta_\iota}{\zeta_\iota^2-\eta_\iota}\frac{1}{p(1-p)}+C^\top_{\iota t}S_{\iota t}C_{2,\iota t}\right)(\overline{A}_{i\iota t}-p)-C^\top_{\iota t}S_{\iota t}(O_{i\iota t}-\mathbb{E}O_{i\iota t})\right\}u_{i\iota t}$$
$$+\sum_{k=t+1}^{M}\left\{S_{\iota k}(O_{i\iota k}-\mathbb{E}O_{i\iota k})-S_{\iota k}C_{1,\iota t}(A_{i\iota k}-p)-S_{\iota k}C_{2,\iota t}(\overline{A}_{i\iota k}-p)\right\}^\top v_{i\iota kt}.$$

Then we have

$$\mathrm{MSE}\left(\widehat{\tau}^c\right)-\sigma^2_{OLS}\asymp\sum_{t=1}^{M}\mathbb{E}\left(\sum_{j_c=1}^{m}\sum_{\iota\in\partial\mathcal{C}_{j_c}}M^c_{1,\iota t}u_{i\iota t}+\sum_{j_c=1}^{m}\sum_{\iota\in\mathcal{C}^o_{j_c}}M^g_{1,\iota t}u_{i\iota t}\right)^2$$
$$+\sum_{t=1}^{M}\sum_{k=t+1}^{M}\mathbb{E}\left\{\sum_{j_c=1}^{m}\sum_{\iota\in\partial\mathcal{C}_{j_c}}(M^i_{2,\iota k})^\top v_{i\iota k}+\sum_{j_c=1}^{m}\sum_{\iota\in\mathcal{C}^o_{j_c}}(M^g_{2,\iota k})^\top v_{i\iota k}\right\}^2,$$

where

$$M^c_{1,\iota t}=\left(\frac{\zeta_\iota-\eta_\iota}{\zeta_\iota^2-\eta_\iota}\frac{1}{p(1-p)}+C^\top_{\iota t}S_{\iota t}C_{1,\iota t}\right)(A^{(j_c)}_{it}-p)-C^\top_{\iota t}S_{\iota t}\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Gamma_{\iota j}(A^{(j_c)}_{it}-p)$$
$$+\left(\frac{1-\eta_\iota}{\zeta_\iota^2-\eta_\iota}\frac{1}{p(1-p)}+C^\top_{\iota t}S_{\iota t}C_{2,\iota t}\right)(\overline{A}_{i\iota t}-p)-C^\top_{\iota t}S_{\iota t}\sum_{j=1}^{t-1}\mathbb{B}_{\iota jt}\Theta_{\iota j}(\overline{A}_{i\iota j}-p).$$

4. Results in temporal fixed, randomized and switchback designs.

The temporal fixed, randomized and switchback designs correspond to $a_{tt'}=1,0$ and $(-1)^{t-t'}$, respectively. In these three designs, we have $a_{jj'}-a_{jt}a_{j't}=0$, which gives

$$N\{\text{MSE}\,(\widehat{\tau}^{c})-\sigma_{OLS}^{2}\}\asymp\sum_{t=1}^{M}\sum_{\iota,\iota'=1}^{R}\frac{1}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t},$$

$$N\{\text{MSE}\,(\widehat{\tau}^{i})-\sigma_{OLS}^{2}\}\asymp\sum_{t=1}^{M}\sum_{\iota=1}^{R}\frac{1}{p(1-p)}\mathbb{E}u_{i\iota t}^{2}+2\sum_{t=1}^{M}\sum_{\substack{\iota=1\\ \iota'\in\mathcal{N}_\iota}}^{R}\frac{n_\iota}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t}$$

$$+\sum_{t=1}^{M}\sum_{\mathcal{N}_\iota\cap\mathcal{N}_\iota'\neq\emptyset}\frac{n_\iota}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t},$$

$$N\{\text{MSE}\,(\widehat{\tau}^{c})-\sigma_{OLS}^{2}\}\asymp\sum_{t=1}^{M}\sum_{j=1}^{m}\frac{1}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t}+2\sum_{t=1}^{M}\sum_{j=1}^{m}\sum_{\iota\in\mathcal{C}_j}\sum_{\iota'\in\mathcal{N}_{\mathcal{C}_j}}\frac{\omega_{\iota'}n_{\iota'}^{(j)}}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t}$$

$$+\sum_{t=1}^{M}\sum_{j=1}^{m}\sum_{\iota,\iota'\in\mathcal{N}_{\mathcal{C}_j}}\frac{\omega_\iota\omega_{\iota'}n_\iota^{(j)}n_{\iota'}^{(j)}}{p(1-p)}\mathbb{E}u_{i\iota t}u_{i\iota' t}.$$

This completes the proof.

3.7. *Proof of Theorem 3.4.* Denote $\mathcal{J}_{\iota,t}=(O_{\iota t},A_{\iota a},m_\iota(\boldsymbol{O}_{\iota t},\boldsymbol{A}_{\iota t}))$ and $\bar{m}_\iota(\boldsymbol{A}_t)=|\mathcal{N}_\iota|^{-1}\sum_{k\in\mathcal{N}_\iota}A_{kt}$. According to Liu et al. (2018) and Uehara, Huang and Jiang (2020), $\|\widehat{\mu}_{\iota t}^{a,(k)}-\mu_{\iota t}^{a}\|_2\|\widehat{Q}_{\iota t}^{a,(k)}-Q_{\iota t}^{a}\|_2=o_p(N^{-1/2})$, $\|\widehat{\mu}_{\iota t}^{a,(k)}-\mu_{\iota t}^{a}\|_2=o_p(1)$, and $\|\widehat{Q}_{\iota t}^{a,(k)}-Q_{\iota t}^{a}\|_2=o_p(1)$ hold for $1\le\iota\le R, 1\le t\le M$, and $1\le k\le K$. Under the global randomized design,

$$\eta_{\iota t}^{1}=\frac{1\{A_t=1\}}{p\{A_t=1\}+(1-p)\{A_t=0\}}=\frac{1}{p}\{A_t=1\}.$$

Under the spatially randomized design,

$$\eta_{\iota t}^{1}=\frac{1}{p^{n_\iota+1}}\{A_{\iota t}=1,\bar{m}_\iota(\boldsymbol{A}_t)=1\}.$$

Note that in both designs, $\{\eta_{jk}\}_{1\le k\le t}$ is independent of $\boldsymbol{O}_t$. From p.45 of the supplement of Kallus and Uehara (2020), $\widehat{\tau}_{DRL}$ attains the following semiparametric efficiency bound

$$\text{Var}\Bigg[\sum_{\iota=1}^{R}(Q_{\iota 1}(\mathbf{1}_{RM})-Q_{\iota 1}(\mathbf{0}_{RM}))$$

$$+\sum_{t=1}^{T}\sum_{\iota=1}^{R}\{\mu_{\iota t}^{1}(Y_{\iota t}^{1}+Q_{j,t+1}^{1}(X_{\iota,t+1}^{1})-Q_{\iota t}^{1}(X_{\iota t}))-\mu_{\iota t}^{0}(Y_{\iota t}^{0}+Q_{j,t+1}^{0}(X_{\iota,t+1}^{0})-Q_{\iota t}^{0}(X_{\iota t}))\}\Bigg]$$

$$=\text{Var}\left\{\sum_{\iota=1}^{R}(Q_{\iota 1}(\mathbf{1}_{RM})-Q_{\iota 1}(\mathbf{0}_{RM}))\right\}$$

$$+\sum_{t=0}^{T}\mathbb{E}\left(\text{Var}\left[\sum_{\iota=1}^{R}\{\mu_{\iota t}^{1}(Y_{\iota t}^{1}+Q_{j,t+1}^{1}(X_{\iota,t+1}^{1})-Q_{\iota t}^{1}(X_{\iota t}))-\mu_{\iota t}^{0}(Y_{\iota t}^{0}+Q_{j,t+1}^{0}(X_{\iota,t+1}^{0})-Q_{\iota t}^{0}(X_{\iota t}))\}\Big|\mathcal{J}_{\iota,t}\right]\right).$$

The second term equals

$$
\begin{aligned}
&\sum_{t=0}^{T}\mathbb{E}\left(\mathrm{Var}\left[\sum_{\iota=1}^{R}\left\{(\mu_{\iota t}^{1}-\mu_{\iota t}^{0})(Y_{\iota t}+Q_{j,t+1}(X_{\iota,t+1})-Q_{\iota t}(X_{\iota t}))\right\}\Big|\mathcal{J}_{\iota,t}\right]\right)\\
=&\sum_{t=0}^{T}\mathbb{E}\left\{\sum_{j_1,j_2}(\mu_{j_1,t}^{1}-\mu_{j_1,t}^{0})(\mu_{j_2,t}^{1}-\mu_{j_2,t}^{0})C_{j_1,j_2,t}\right\}\\
=&\sum_{t=1}^{T}\mathbb{E}\left\{\sum_{j_1,j_2}(\eta_{j_1,t}^{1}-\eta_{j_1,t}^{0})(\eta_{j_2,t}^{1}-\eta_{j_2,t}^{0})\right\}\widetilde{C}_{j_1,j_2,t},
\end{aligned}
$$

where $\widetilde{C}_{j_1,j_2,t}=\omega_{j_1}\omega_{j_2}C_{j_1,j_2,t}$. where the last equation is obtained by the fact that the design is independent of the measurement errors. Under the global randomized design,

$$
\begin{aligned}
&\mathbb{E}\left\{\sum_{j_1,j_2}(\eta_{j_1,t}^{1}-\eta_{j_1,t}^{0})(\eta_{j_2,t}^{1}-\eta_{j_2,t}^{0})\right\}\widetilde{C}_{j_1,j_2,t}\\
=&\mathbb{E}\left\{\sum_{j_1,j_2}\left(\frac{1}{p}\{a_t=1\}-\frac{1}{1-p}\{a_t=0\}\right)^2\right\}\widetilde{C}_{j_1,j_2,t}\\
=&\left(\frac{1}{p}+\frac{1}{1-p}\right)\sum_{j_1,j_2}\widetilde{C}_{j_1,j_2,t}.
\end{aligned}
$$

For the spatio-random design,

$$
\begin{aligned}
&\mathbb{E}\left\{\sum_{j_1,j_2}(\eta_{j_1,t}^{1}-\eta_{j_1,t}^{0})(\eta_{j_2,t}^{1}-\eta_{j_2,t}^{0})\right\}\widetilde{C}_{j_1,j_2,t}\\
=&\mathbb{E}\left\{\sum_{j_1,j_2}\left(\frac{1}{p^{n_{j_1}+1}}\{a_{j_1,t}=1,\bar{m}_{j_1}(\boldsymbol{a}_t)=1\}-\frac{1}{(1-p)^{n_{j_1}+1}}\{a_{j_1,t}=0,\bar{m}_{j_1}(\boldsymbol{a}_t)=0\}\right)\right.\\
&\left.\left(\frac{1}{p^{n_{j_2}+1}}\{a_{j_2,t}=1,\bar{m}_{j_2}(\boldsymbol{a}_t)=1\}-\frac{1}{(1+p)^{n_{j_2}+1}}\{a_{j_2,t}=0,\bar{m}_{j_2}(\boldsymbol{a}_t)=0\}\right)\right\}\widetilde{C}_{j_1,j_2,t}\\
=&\sum_{j_1\neq j_2}\mathbb{E}\left(\frac{1}{p^{n_{j_1}+1}}\{a_{j_1,t}=1,\bar{m}_{j_1}(\boldsymbol{a}_t)=1\}-\frac{1}{(1-p)^{n_{j_1}+1}}\{a_{j_1,t}=0,\bar{m}_{j_1}(\boldsymbol{a}_t)=0\}\right)\\
&\cdot\mathbb{E}\left(\frac{1}{p^{n_{j_2}+1}}\{a_{j_2,t}=1,\bar{m}_{j_2}(\boldsymbol{a}_t)=1\}-\frac{1}{(1-p)^{n_{j_2}+1}}\{a_{j_2,t}=0,\bar{m}_{j_2}(\boldsymbol{a}_t)=0\}\right)\widetilde{C}_{j_1,j_2,t}\\
&+\sum_{\iota=1}^{R}\mathbb{E}\left(\frac{1}{p^{n_\iota+1}}\{a_{j,t}=1,\bar{m}_{\iota}(\boldsymbol{a}_t)=1\}-\frac{1}{(1-p)^{n_\iota+1}}\{a_{j,t}=0,\bar{m}_{\iota}(\boldsymbol{a}_t)=0\}\right)^2\widetilde{C}_{\iota,t}\\
=&\sum_{\iota=1}^{R}\left(\frac{1}{p^{n_\iota+1}}+\frac{1}{(1-p)^{n_\iota+1}}\right)\widetilde{C}_{\iota,t}+\sum_{\substack{\iota\neq\iota'\\ \mathcal{N}_\iota\cap\mathcal{N}_{\iota'}\neq\emptyset}}\left(\frac{1}{p^{m_{\iota\iota'}}}+\frac{1}{(1-p)^{m_{\iota\iota'}}}\right)\widetilde{C}_{\iota,\iota',t}.
\end{aligned}
$$

This completes the proof.

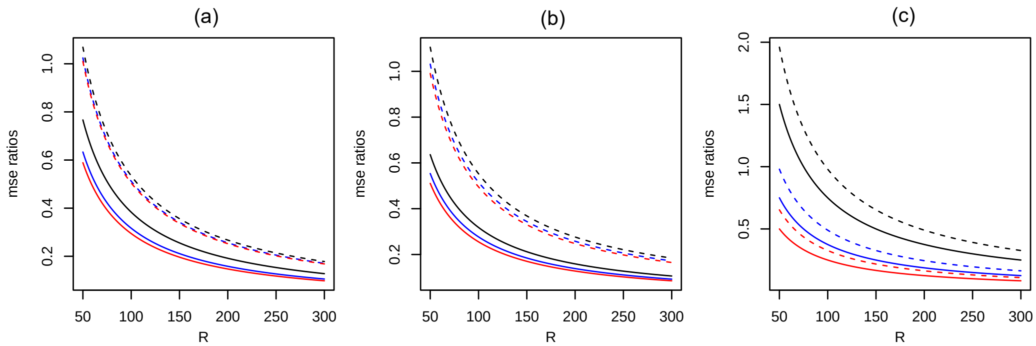

Fig 1: The figure illustrates the MSE ratios $\mathrm{MSE}(\widehat{\tau}^i)/\mathrm{MSE}(\widehat{\tau}^g)$ (solid lines) and $\mathrm{MSE}(\widehat{\tau}^c)/\mathrm{MSE}(\widehat{\tau}^g)$ with $c=9$ (dashed lines) across three distinct scenarios: Panel (a) corresponds to Example 1, Panel (b) to Example 2, and Panel (c) to Example 2. The urban area is segmented into uniform squares with $r=4$, and the total number of regions R spans from 50 to 300. The black, blue, and red lines represent the MSE ratios for hyperparameter ρ values of $0.3, 0.6$, and 0.9, respectively, as detailed in Examples 1-3.

4. Additional Noise Structures and Simulations.

4.1. *Typical examples.* We first use three examples to show that spatially randomized designs generally yield estimators with lower MSEs compared to the global design in the nondynamic settings; see also Figure 1 for numerical comparisons. The superiority of the spatially randomized designs becomes increasingly significant with large R.

EXAMPLE. (Exchangeable) There exists some constant $0<\rho<1$ such that $\mathrm{cov}(e_\iota, e_{\iota'}) = \mathbb{I}(\iota=\iota') + \rho\mathbb{I}(\iota\neq\iota')$. Then, we have $\mathrm{MSE}\left(\widehat{\tau}^i\right)/\mathrm{MSE}\left(\widehat{\tau}^g\right) \lesssim (r+1)^2/[1+\rho(R-1)]$.

EXAMPLE. (Exponential-decay) Let $(x_\iota, y_\iota)\in(0,1)^2$ denote the coordinates of the center of the ιth region after scaling. Additionally, let $d_{\iota\iota'} = \left\{(x_\iota - x_{\iota'})^2 + (y_\iota - y_{\iota'})^2\right\}^{1/2}/2$ denote the spatial distance between (x_ι, y_ι) and $(x_{\iota'}, y_{\iota'})$. Assume there exists some $0<\rho<1$ such that $\mathrm{cov}(e_\iota, e_{\iota'}) = \rho^{d_{\iota\iota'}}$. Then, there exists some $0<\delta\leq\sqrt{2}$ such that $\mathrm{MSE}\left(\widehat{\tau}^i\right)/\mathrm{MSE}\left(\widehat{\tau}^g\right) \lesssim (1+r)^2/[1+(R-1)\rho^\delta]$.

EXAMPLE. (d-dependent) There exist some constant $0<\rho<1$ such that $\mathrm{cov}(e_\iota, e_{\iota'}) = \mathbb{I}(\iota=\iota') + (\rho - \frac{|\iota-\iota'|}{R})\mathbb{I}(|\iota-\iota'|\leq\rho R)$. Then, we have $\mathrm{MSE}\left(\widehat{\tau}^i\right)/\mathrm{MSE}\left(\widehat{\tau}^g\right) \lesssim (1+r)^2/[1+(1-\rho)\rho^2 R]$.

To see why the above conclusions hold, it suffices to provide a lower bound for ν to obtain an upper bound for the ratio $\mathrm{MSE}\left(\widehat{\tau}^i\right)/\mathrm{MSE}\left(\widehat{\tau}^g\right)$. Then

1. In the exchangeable case, we have

$$\nu = \frac{R + R(R-1)\rho}{R} = 1 + (R-1)\rho.$$

2. In the exponential-decay case, we have

$$\nu = \frac{R + \sum_{\iota\neq\iota'}\rho^{d_{\iota\iota'}}}{R} \geq \frac{R + R(R-1)\rho^\delta}{R} = 1 + (R-1)\rho^\delta,$$

where $\delta = \max_{\iota,\iota'} d_{\iota\iota'}$.

3. Note that $\#\{(\iota,\iota') : |\iota-\iota'| = k\} = 2(R-k)$. In the d-dependent case, we have

$$\nu = \frac{R+\sum_{\iota\neq\iota'}\mathbb{V}_{\iota\iota'}}{R} = \frac{R+\sum_{k=1}^{\rho R}2(R-k)\cdot(\rho-k/R)}{R} \geq 1+2(1-\rho)\sum_{k=1}^{\rho R}(\rho-k/R)$$
$$= 1+(1-\rho)\rho^2 R.$$

4.2. *Simulation results.* Then we present more estimation results of single-stage case. The coefficients and spatial configurations are the same as Section 4.1 in the main context. The empirical MSE ratios in the parametric regression and semiparametric regression when there is no interference with the noise taking the covariance structure of Example 1–3 are shown in Tables 1 and 2. Results for interference-existing case are shown in Tables 3 and 4.

TABLE 1
Empirical values of $r_1 = \mathrm{MSE}(\widehat{\tau}^i)/\mathrm{MSE}(\widehat{\tau}^g)$ *and* $r_2 = \mathrm{MSE}(\widehat{\tau}^c)/\mathrm{MSE}(\widehat{\tau}^g)$ *in the parametric regression of nondynamic stage without interference.*

			Example 1			Example 2			Example 3		
r	ρ	ratio	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$
3	0.9	r_1	0.031	0.016	0.009	0.028	0.015	0.008	0.045	0.020	0.013
		r_2	0.229	0.128	0.062	0.229	0.129	0.061	0.259	0.140	0.073
	0.6	r_1	0.049	0.020	0.014	0.032	0.015	0.010	0.100	0.043	0.026
		r_2	0.254	0.126	0.070	0.253	0.132	0.068	0.408	0.215	0.117
	0.3	r_1	0.097	0.038	0.028	0.040	0.016	0.012	0.340	0.156	0.098
		r_2	0.295	0.132	0.082	0.288	0.142	0.079	0.826	0.419	0.256
4	0.9	r_1	0.031	0.016	0.009	0.028	0.015	0.008	0.045	0.020	0.013
		r_2	0.229	0.128	0.062	0.229	0.129	0.061	0.259	0.140	0.073
	0.6	r_1	0.049	0.020	0.014	0.032	0.015	0.010	0.100	0.043	0.026
		r_2	0.254	0.126	0.070	0.253	0.132	0.068	0.408	0.215	0.117
	0.3	r_1	0.097	0.038	0.028	0.040	0.016	0.012	0.340	0.156	0.098
		r_2	0.295	0.132	0.082	0.288	0.142	0.079	0.826	0.419	0.256
6	0.9	r_1	0.031	0.016	0.009	0.028	0.015	0.008	0.045	0.020	0.013
		r_2	0.229	0.128	0.062	0.229	0.129	0.061	0.259	0.140	0.073
	0.6	r_1	0.049	0.020	0.014	0.032	0.015	0.010	0.100	0.043	0.026
		r_2	0.254	0.126	0.070	0.253	0.132	0.068	0.408	0.215	0.117
	0.3	r_1	0.097	0.038	0.028	0.040	0.016	0.012	0.340	0.156	0.098
		r_2	0.295	0.132	0.082	0.288	0.142	0.079	0.826	0.419	0.256

For the inference performances, we present the following results:

(1) Note that the inference results of the parametric regression and semiparametric regression when there exists interference with noise covariance in Example 2 are plotted in Figure 3 of the main context. We show the corresponding results when there is no interference in Figure 2 below.
(2) Then we present the inference results of the parametric regression when the noise takes the covariance structure of Example 1 and 3 in Figure 3–4.
(3) Finally we present the inference results of the semiparametric regression when the noise takes the covariance structure of Example 1 and 3 in Figure 3–4.

TABLE 2
Empirical values of $r_1 = \mathrm{MSE}(\widehat{\tau}^i_{DR})/\mathrm{MSE}(\widehat{\tau}^g_{DR})$ *and* $r_2 = \mathrm{MSE}(\widehat{\tau}^c_{DR})/\mathrm{MSE}(\widehat{\tau}^g_{DR})$ *in the nonparametric regression of nondynamic stage without interference.*

			Example 1			Example 2			Example 3		
r	ρ	ratio	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$
3	0.9	r_1	0.028	0.013	0.008	0.026	0.012	0.008	0.040	0.017	0.011
		r_2	0.126	0.072	0.044	0.122	0.075	0.042	0.244	0.107	0.064
	0.6	r_1	0.044	0.018	0.012	0.030	0.013	0.009	0.089	0.036	0.022
		r_2	0.148	0.077	0.049	0.132	0.080	0.045	0.384	0.171	0.102
	0.3	r_1	0.093	0.037	0.023	0.038	0.016	0.010	0.290	0.126	0.077
		r_2	0.188	0.097	0.060	0.150	0.089	0.050	0.789	0.350	0.240
4	0.9	r_1	0.035	0.014	0.008	0.033	0.013	0.008	0.051	0.019	0.011
		r_2	0.288	0.114	0.061	0.276	0.116	0.059	0.334	0.119	0.064
	0.6	r_1	0.054	0.020	0.012	0.037	0.014	0.009	0.113	0.040	0.023
		r_2	0.390	0.128	0.069	0.310	0.124	0.064	0.581	0.199	0.104
	0.3	r_1	0.113	0.040	0.023	0.047	0.017	0.010	0.359	0.136	0.081
		r_2	0.675	0.183	0.092	0.378	0.141	0.074	1.477	0.445	0.251
6	0.9	r_1	0.029	0.013	0.008	0.027	0.013	0.008	0.041	0.018	0.011
		r_2	0.203	0.100	0.054	0.198	0.104	0.052	0.265	0.110	0.063
	0.6	r_1	0.045	0.018	0.012	0.031	0.013	0.008	0.091	0.038	0.022
		r_2	0.251	0.104	0.061	0.219	0.112	0.058	0.426	0.178	0.100
	0.3	r_1	0.094	0.038	0.023	0.039	0.016	0.010	0.293	0.128	0.078
		r_2	0.368	0.129	0.076	0.259	0.127	0.068	0.945	0.375	0.237

TABLE 3
Empirical values of $r_1 = \mathrm{MSE}(\widehat{\tau}^i)/\mathrm{MSE}(\widehat{\tau}^g)$ *and* $r_2 = \mathrm{MSE}(\widehat{\tau}^c)/\mathrm{MSE}(\widehat{\tau}^g)$ *in the parametric regression of nondynamic stage with interference.*

			Example 1			Example 2			Example 3		
r	ρ	ratio	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$
3	0.9	r_1	0.366	0.211	0.121	0.360	0.214	0.117	0.420	0.226	0.142
		r_2	0.525	0.367	0.188	0.522	0.370	0.185	0.572	0.396	0.210
	0.6	r_1	0.417	0.204	0.151	0.379	0.200	0.134	0.693	0.356	0.238
		r_2	0.569	0.347	0.203	0.554	0.361	0.196	0.817	0.531	0.294
	0.3	r_1	0.556	0.244	0.208	0.421	0.195	0.155	1.454	0.846	0.625
		r_2	0.626	0.340	0.222	0.593	0.361	0.212	1.101	0.873	0.588
4	0.9	r_1	0.552	0.308	0.191	0.544	0.316	0.187	0.658	0.321	0.227
		r_2	0.641	0.456	0.250	0.641	0.457	0.246	0.723	0.469	0.286
	0.6	r_1	0.623	0.300	0.232	0.582	0.306	0.219	1.015	0.501	0.374
		r_2	0.694	0.444	0.273	0.690	0.460	0.272	0.956	0.663	0.406
	0.3	r_1	0.815	0.365	0.302	0.661	0.319	0.259	2.167	1.191	0.883
		r_2	0.762	0.454	0.300	0.759	0.485	0.309	1.306	0.984	0.64
6	0.9	r_1	1.006	0.562	0.358	0.982	0.577	0.357	1.182	0.585	0.415
		r_2	0.675	0.491	0.274	0.674	0.493	0.271	0.762	0.510	0.316
	0.6	r_1	1.120	0.530	0.413	1.042	0.557	0.413	1.779	0.881	0.653
		r_2	0.729	0.478	0.300	0.724	0.497	0.298	0.98	0.708	0.448
	0.3	r_1	1.395	0.594	0.499	1.171	0.577	0.485	3.512	1.877	1.339
		r_2	0.801	0.488	0.326	0.795	0.524	0.339	1.308	1.014	0.667

TABLE 4
Empirical values of $r_1 = \mathrm{MSE}(\hat{\tau}^i_{DR})/\mathrm{MSE}(\hat{\tau}^g_{DR})$ *and* $r_2 = \mathrm{MSE}(\hat{\tau}^c_{DR})/\mathrm{MSE}(\hat{\tau}^g_{DR})$ *in the nonparametric regression of nondynamic stage with interference.*

			Example 1			Example 2			Example 3		
r	ρ	ratio	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$	$R=36$	$R=64$	$R=144$
3	0.9	r_1	0.188	0.103	0.065	0.186	0.103	0.062	0.223	0.125	0.078
		r_2	0.315	0.156	0.092	0.306	0.163	0.089	0.359	0.170	0.108
	0.6	r_1	0.236	0.116	0.079	0.202	0.106	0.068	0.403	0.212	0.138
		r_2	0.363	0.160	0.101	0.331	0.173	0.095	0.533	0.258	0.160
	0.3	r_1	0.379	0.175	0.111	0.239	0.118	0.079	0.975	0.550	0.357
		r_2	0.448	0.189	0.119	0.378	0.192	0.107	0.874	0.479	0.343
4	0.9	r_1	0.362	0.170	0.100	0.352	0.165	0.096	0.482	0.214	0.123
		r_2	0.452	0.192	0.105	0.438	0.199	0.102	0.548	0.212	0.122
	0.6	r_1	0.465	0.204	0.122	0.382	0.175	0.107	0.854	0.373	0.212
		r_2	0.564	0.204	0.116	0.484	0.214	0.111	0.881	0.339	0.184
	0.3	r_1	0.799	0.316	0.172	0.459	0.204	0.127	2.212	0.898	0.539
		r_2	0.843	0.259	0.139	0.576	0.244	0.129	2.052	0.717	0.393
6	0.9	r_1	0.297	0.141	0.083	0.286	0.138	0.080	0.380	0.174	0.103
		r_2	0.373	0.182	0.105	0.363	0.190	0.102	0.435	0.198	0.121
	0.6	r_1	0.381	0.172	0.101	0.316	0.153	0.090	0.682	0.290	0.181
		r_2	0.440	0.188	0.115	0.399	0.203	0.111	0.627	0.303	0.179
	0.3	r_1	0.625	0.265	0.141	0.385	0.185	0.107	1.714	0.742	0.493
		r_2	0.579	0.225	0.135	0.466	0.231	0.129	1.166	0.598	0.373

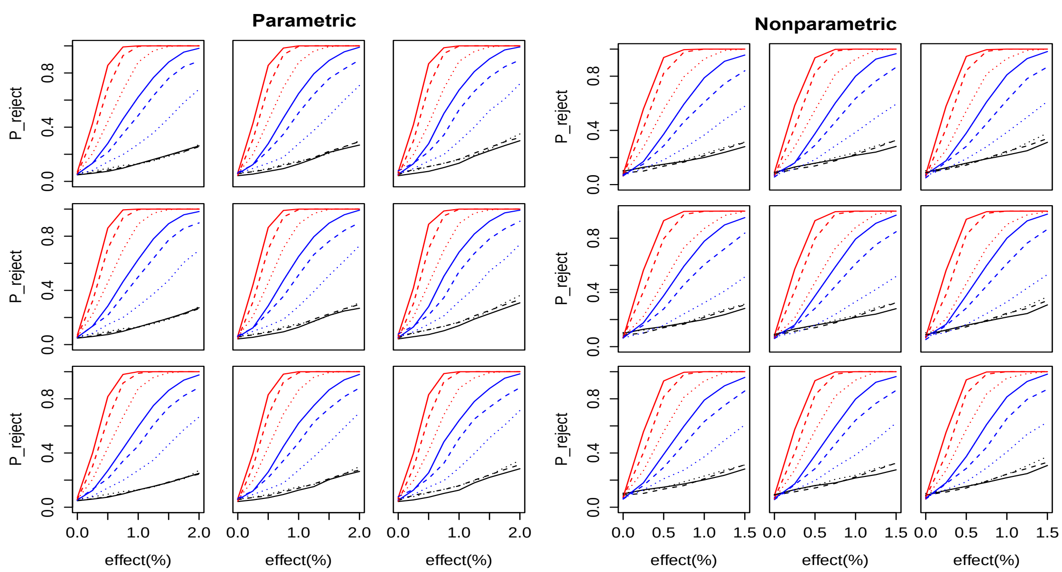

Fig 2: Rejection probability in the nonparametric regression of nondynamic setting under different relative improvements of the new policy. The red, blue and black lines represent to the individual-, cluster and global-randomized designs, with $R = 144, 81, 36$ in solid, dashed and dotted lines, respectively. The three rows of panels correspond to $r = 6, 4, 3$, and the three columns correspond to $\rho = 0.9, 0.6, 0.3$, respectively.

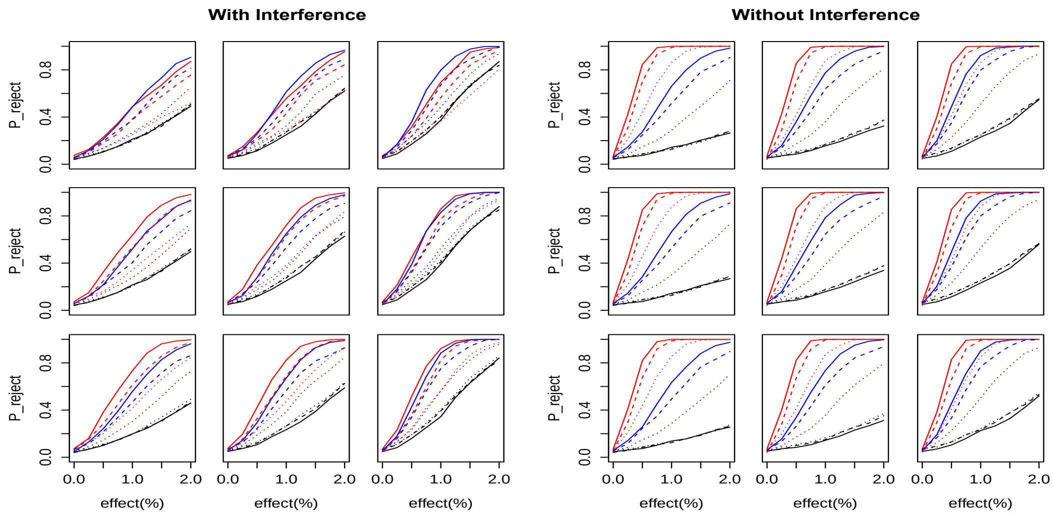

Fig 3: Rejection probability in the parametric regression of single-stage case with and without interference when the noise takes the covariance structure of Example 1, and the relative improvement of the new policy ranges from 0% to 2.0%. The red lines represents the individual-randomized design and the blues ones are the results of the cluster-randomized design, with the black ones corresponding to the global randomized design. The number of regions $R = 144, 81, 36$ are plotted in solid, dashed and dotted lines, respectively. The three rows of panels correspond to $r = 6, 4, 3$, and the three columns correspond to $\rho = 0.9, 0.6, 0.3$, respectively.

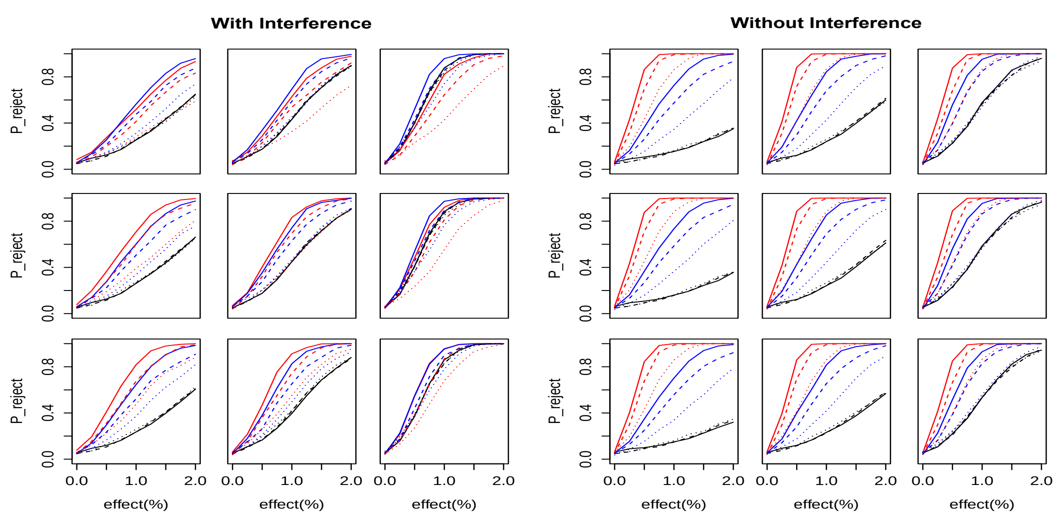

Fig 4: Rejection probability in the parametric regression of single-stage case with and without interference when the noise takes the covariance structure of Example 3, and the relative improvement of the new policy ranges from 0% to 2.0%. The red lines represents the individual-randomized design and the blues ones are the results of the cluster-randomized design, with the black ones corresponding to the global randomized design. The number of regions $R = 144, 81, 36$ are plotted in solid, dashed and dotted lines, respectively. The three rows of panels correspond to $r = 6, 4, 3$, and the three columns correspond to $\rho = 0.9, 0.6, 0.3$, respectively.

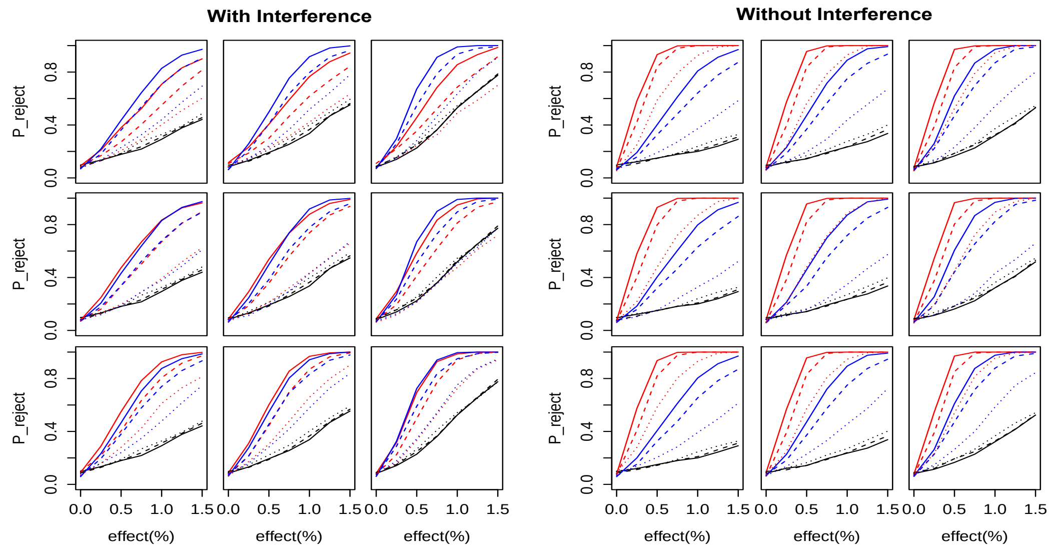

Fig 5: Rejection probability in the nonparametric regression of single-stage case when the relative improvement of the new policy is $0\%, 0.25\%, 0.5\%, 0.75\%, 1.0\%, 1.25\%$ and 1.5%, and the noise takes the covariance structure of Example 1. The red, blue and black lines represent to the individual-, cluster and global-randomized designs, with $R = 144, 81, 36$ in solid, dashed and dotted lines, respectively. The three rows of panels correspond to $r = 6, 4, 3$, and the three columns correspond to $\rho = 0.9, 0.6, 0.3$, respectively.

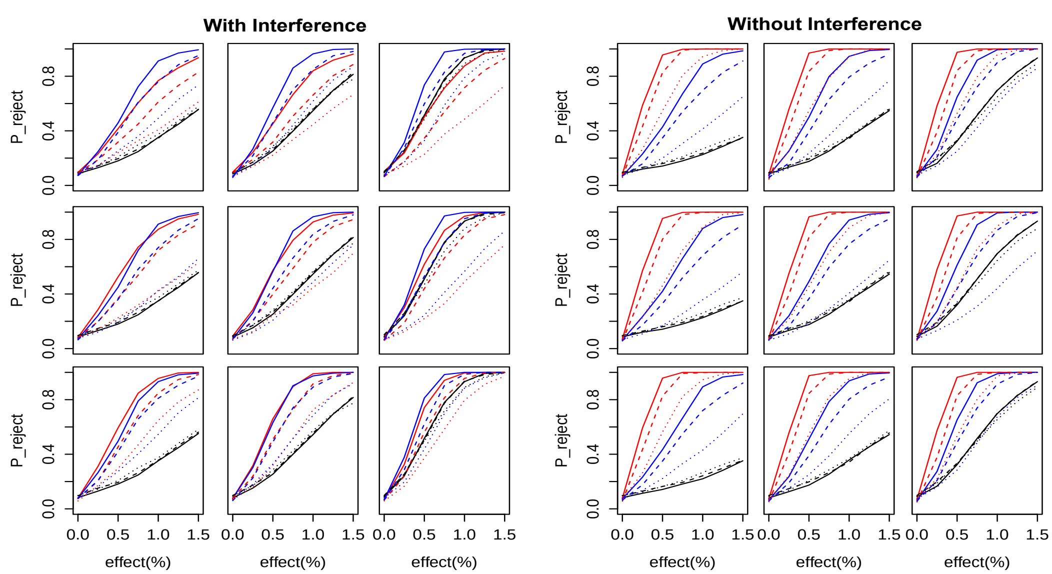

Fig 6: Rejection probability in the nonparametric regression of single-stage case when the relative improvement of the new policy is $0\%, 0.25\%, 0.5\%, 0.75\%, 1.0\%, 1.25\%$ and 1.5%, and the noise takes the covariance structure of Example 3. The red, blue and black lines represent to the individual-, cluster and global-randomized designs, with $R = 144, 81, 36$ in solid, dashed and dotted lines, respectively. The three rows of panels correspond to $r = 6, 4, 3$, and the three columns correspond to $\rho = 0.9, 0.6, 0.3$, respectively.

REFERENCES

KALLUS, N. and UEHARA, M. (2020). Double reinforcement learning for efficient off-policy evaluation in markov decision processes. *Journal of Machine Learning Research* **21** 6742-6804.

LIU, Q., LI, L., TANG, Z. and ZHOU, D. (2018). Breaking the curse of horizon: Infinite-horizon off-policy estimation. *Advances in Neural Information Processing Systems* **31**.

SHI, C., ZHU, J., YE, S., LUO, S., ZHU, H. and SONG, R. (2022). Off-policy confidence interval estimation with confounded markov decision process. *Journal of the American Statistical Association* in press.

SHI, C., WAN, R., SONG, G., LUO, S., ZHU, H. and SONG, R. (2023). A multiagent reinforcement learning framework for off-policy evaluation in two-sided markets. *The Annals of Applied Statistics* **17** 2701–2722.

UEHARA, M., HUANG, J. and JIANG, N. (2020). Minimax weight and q-function learning for off-policy evaluation. In *International Conference on Machine Learning* 9659–9668. PMLR.