

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*

This article studies the benefits of using spatially randomized experimental designs which partition the experimental area into non-overlapping units with treatments assigned randomly to these units. Such designs improve policy evaluation in online experiments by providing more precise policy value estimators and more effective testing algorithms than traditional global designs, which apply the same treatment across all units simultaneously. We examine both parametric and nonparametric methods for estimating and inferring policy values based on the spatially randomized designs. Our analysis includes evaluating the mean squared error of the treatment effect estimator and the statistical power of the associated tests. Additionally, we extend our findings to the dynamic setting with spatio-temporal dependencies, where treatments are allocated sequentially over time, and account for potential temporal carryover effects. Our theoretical insights are supported by comprehensive numerical experiments.

1. Introduction. Policy evaluation in spatially dependent experiments involves analyzing spatially referenced data to assess the impact of new products. This methodology is widely used in diverse fields, including environmental studies (Zigler, Dominici and Wang, 2012), epidemiology (Hudgens and Halloran, 2008; Callaway and Li, 2023), social science (Sobel, 2006), and technology industries (Zhou et al., 2024; Wang et al., 2025). In such experiments, the challenge often lies in the limited number of observations and the small magnitude of treatment effects. Moreover, the strong interconnections among spatio-temporal units tend to increase the variance of estimations, making it difficult to detect the weak effects. Additionally, the treatment on one spatial unit may influence the outcomes of other units, leading to a breach of the stable unit treatment value assumption (SUTVA, see, e.g., Imbens and Rubin, 2015), and causing interference or spillover effects that complicate the analysis (Basse et al., 2024).

As an illustration, consider the applications in ride-sourcing platforms such as Uber, Lyft, and Didi Chuxing. These companies extensively utilize A/B testing to assess the efficacy of universal treatment policies, such as new order dispatch or subsidy strategies implemented across an entire city (Xu et al., 2018; Zhou et al., 2021; Luo et al., 2024). In these settings, the dynamic networks of call orders and available drivers represent the supply and demand within these marketplaces, exhibiting significant spatial correlations (Ke et al., 2018). Moreover, budget limitations often restrict the duration of online experiments to a mere two weeks

Keywords and phrases: A/B testing, policy evaluation, reinforcement learning, spatially randomized designs.

(Shi et al., 2023a), with the anticipated improvements from novel policies being relatively modest, typically between 0.5% and 2% (Tang et al., 2019). To understand the spillover effects, consider that implementing a subsidy policy for drivers in one area might draw drivers from adjacent units, thus influencing outcomes in those neighboring areas as well. This interconnected structure highlights the limitations of conventional experimental designs and calls for novel frameworks tailored to such spatially and temporally dependent systems.

Experimental design plays a crucial role in accurately estimating and inferring causal effects, particularly in complex settings (Seltman, 2012; Imai, Tingley and Yamamoto, 2013; Jia, Kallus and Yu, 2023; Li et al., 2023; Ding, 2024; Sun et al., 2024; Wen et al., 2025; Wu et al., 2026). In clinical trials, numerous adaptive designs have been developed to personalize treatments based on individual-level information (see e.g., Hu and Hu, 2012; Liu and Hu, 2022; Ma et al., 2024). However, these methods typically assume that observations are independent and identically distributed (i.i.d.), limiting their applicability in settings with spatial or temporal dependencies. Recently, there has been growing interest in experimental designs that account for spatial dependence and interference (Jagadeesan, Pillai and Volfovsky, 2020; Kong, Yuan and Zheng, 2021; Leung, 2022; Zhou et al., 2024; Zhu et al., 2025). However, these works generally assume independent noise across spatial units and ignore dynamic settings that capture temporal carryover effects, which are unique features in two-sides markets such as ridesharing companies. Consequently, existing literature provides limited guidance on how to design experiments in the presence of both spatial dependence and temporal dynamics, leaving a gap in practical methodologies for modern two-sided marketplaces.

This paper makes the following contributions. First, we formulate a unified framework for analyzing three major designs in spatially randomized experiments: individual-randomized, cluster-randomized, and global designs. The global design, which applies a uniform treatment to all units, serves as a natural baseline for assessing the efficiency of spatially heterogeneous treatment allocations. To accommodate diverse data environments, we study both parametric and semiparametric estimation procedures for evaluating treatment effects. We further extend our framework to dynamic environments characterized by spatio-temporal dependencies, enabling the study of sequential interventions and temporal spillover effects. Second, we establish quantitative results that characterize the mean squared error (MSE) and testing power of the proposed estimators, clarifying how design efficiency depends on spatial correlation, interference range, and cluster size. The analysis also provides insight into how to choose between individual- and cluster-randomized designs, and offers guidance on selecting an optimal cluster size that balances within- and between-cluster dependencies. Third, our framework directly connects to modern A/B testing and policy evaluation practices in large-scale online platforms such as ride-hailing, e-commerce, and digital advertising. In these applications, spatial or network interference naturally arises when localized interventions such as subsidies, pricing, or dispatch strategies affect outcomes in neighboring areas. Our studied designs offer a principled way to account for such spillovers, enhancing both the validity and efficiency of empirical policy evaluations.

1.1. *Related work.* In this subsection, we review existing works that are related to our work as follows. The most prevalent approach for inferring treatment effects in the aforementioned contexts is off-policy evaluation (OPE, see e.g. Uehara, Shi and Kallus, 2026, for a review), which aims to evaluate the impact of a target policy offline using a pre-collected historical dataset generated by a different behavior policy. In the context of finite-horizon settings characterized by a limited number of decision points, (augmented) inverse propensity score weighted estimators have been introduced for policy evaluation (Jiang and Li, 2016; Luedtke and Van Der Laan, 2016; Thomas and Brunskill, 2016; Hanna, Niekum and Stone, 2019; Zhou et al., 2025). More recent advancements have extended these methodologies to

efficiently handle evaluations over extended or infinite time horizons (Shi et al., 2021, 2022; Wang, Qi and Wong, 2023; Kallus and Uehara, 2022; Zhou et al., 2023a). However, these studies have not explored policy evaluation in the context of spatial interference or the role of experimental design in such settings.

Research on spatial interference has primarily branched into two prominent types of methods. The first type is the partial interference, segmenting individuals into clusters with the interference effects contained within each respective cluster (Liu, Hudgens and Becker-Dreps, 2016; Zigler and Papadogeorgou, 2021; Huber and Steinmayr, 2021). The second type is the local or network interference, where the interference effects are confined to the local network of each unit (Perez-Heydrich et al., 2014; Puelz et al., 2022; Shi et al., 2023b). Recent studies have proposed more complex interference structures to accommodate specific application problems (Aronow and Samii, 2017; Tchetgen, Fulcher and Shpitser, 2021; Larsen et al., 2022; Dai et al., 2024). However, these works did not consider experimental designs, which is the focus of this paper.

1.2. Outline of the paper. The rest of the paper is organized as follows. In Section 2, we present the problem formulation in the nondynamic setting and utilize parametric and semi-parametric methods to estimate the treatment effects. We then examine the MSEs of these estimators under different spatial designs and compare their testing efficiency. Section 3 extends our analysis to the dynamic case. The numerical simulations and real data applications are displayed in Section 4, which further verifies the usefulness of our method. Technical proofs are collected in Sections 2 and 3 of the Supplementary Material (Yang et al., 2026).

2. Nondynamic Setting. We start our exploration in the nondynamic setting, and the examination of the dynamic setting will be presented in the next section.

2.1. Problem formulation. Consider a city divided into R non-overlapping spatial units and we are interested in evaluating the performance of a newly developed policy against a standard control. For the ι th unit, let $Y_\iota(1)$ and $Y_\iota(0)$ represent the potential daily outcomes under the new and existing policies for the whole city, respectively. Our focus is the average treatment effect (ATE),

$$\tau = \sum_{\iota=1}^R \mathbb{E} \{Y_\iota(\mathbf{1}_R) - Y_\iota(\mathbf{0}_R)\},$$

where $Y_\iota(\mathbf{a}_R)$ denotes the potential outcome of unit ι under the global assignment vector $\mathbf{A} = \mathbf{a}_R = (a, \dots, a)^T \in \mathbb{R}^R$ with $a \in \{0, 1\}$. Our definition of ATE aligns with those in Forastiere, Airolidi and Mealli (2021); Leung (2022) and Lu et al. (2023). Consider the following one-sided test hypotheses:

$$H_0 : \tau \leq 0 \text{ V.S. } H_1 : \tau > 0.$$

Under the null hypothesis, the improvement of benefits brought by the new policy is relatively low compared to the implementation costs. As such, we recommend to use the standard control.

To estimate and test the ATE, we collect data over N days in the form of treatment-outcome pairs $\{(A_{i,\iota}, Y_{i,\iota}) : 1 \leq i \leq N, 1 \leq \iota \leq R\}$ where $A_{i,\iota} \in \{0, 1\}$ denotes the binary treatment assignment for unit ι on day i , and $Y_{i,\iota}$ is the corresponding observed outcome. We also incorporate observed covariates $\{O_{i,\iota}\}_{i,\iota}$, such as the number of active drivers in a ride-sharing platform, to improve estimation precision. We assume the consistency assumption (CA) as follows:

–**CA.** For any i and ι , $Y_{i,\iota}$ equals to the potential outcome $Y_{i,\iota}(A_{i,\iota})$.

The aim of this paper is to compare the performance of different experimental designs (in particular, two spatially randomized designs versus a global design) in terms of policy evaluation accuracy. We evaluate their relative performance by analyzing the mean squared error (MSE) of the ATE estimators and the power of hypothesis testing procedures. To facilitate comparison, we assume that observations across days are independent, i.e., daily outcomes are not influenced by treatment assignments from previous days.

In the *global design*, a uniform policy is applied across all R units for each day, i.e. $A_{i,1} = A_{i,2} = \dots = A_{i,R}$ for any given day i . In contrast, spatially randomized designs allow these $A_{i,s}$ to be different at each time. We study two specific types of designs. The first one is the *individual-randomized design* that allocates treatments to each unit independently with a non-zero probability p_ι of receiving treatment 1. The other one is the *cluster-randomized design* which organizes units into m non-overlapping clusters $\{\mathcal{C}_1, \dots, \mathcal{C}_m\}$ based on spatial proximity and ensures uniform treatment within each cluster. Specifically, for any cluster \mathcal{C}_j , we have $A_{i,k_1} = A_{i,k_2} = A_j^{(i)}$ for any $k_1, k_2 \in \mathcal{C}_j$, where $A_j^{(i)}$ is assigned independently with probability $p^{(j)}$. In practice, clusters are often formed by grouping spatially adjacent or strongly interacting units to reduce between-cluster interference. Typical methods include community detection or hierarchical clustering. Our theoretical analysis further suggests that the optimal cluster size should scale with the interference range, emphasizing the importance of accurately capturing the interference structure. Notably, the individual-randomized design is a special cluster-randomized design with $m = R$ and $\mathcal{C}_j = \{j\}$ for every j .

In spatial settings, outcomes may be affected not only by a unit's own treatment but also by other units' treatments, a phenomenon known as interference. We model such effects via an interference function $f_{\theta_\iota}(\{A_{i,j}\}_{j \in \mathcal{N}_\iota})$ where \mathcal{N}_ι represents the interference neighbor set of unit ι . That is, unit $j \in \mathcal{N}_\iota$ if the treatment on unit j affects the outcome of unit ι . These neighborhoods may be heterogeneous across units and can also be unknown and learned from data (Yuan, Altenburger and Kooti, 2021; Zhang, Yang and Yao, 2026). While various forms of $f_{\theta_\iota}(\cdot)$ can be considered such as identity functions, linear combinations, or thresholded rules, we focus on the mean-field approximation: $\bar{A}_{i,\iota} = n_\iota^{-1} \sum_{k \in \mathcal{N}_\iota} A_{ik}$, where $n_\iota = |\mathcal{N}_\iota|$. This formulation effectively summarizes the collective influence of neighboring treatments and has been widely adopted in the literature (Yang et al., 2018; Luo et al., 2024; Hu, Li and Wager, 2022; Shi et al., 2023b). Our results can be extended to other functional forms as discussed in related work.

2.2. Parametric and semiparametric learning. We now introduce the parametric and semiparametric learning methods in the nondynamic setting.

Parametric learning. We introduce the following parametric outcome model,

$$(1) \quad Y_{i,\iota} = \alpha_\iota + O_{i,\iota}^\top \beta_\iota + \gamma_\iota A_{i,\iota} + \theta_\iota \bar{A}_{i,\iota} + e_{i,\iota},$$

where $\alpha_\iota, \gamma_\iota \in \mathbb{R}$ and $\beta_\iota \in \mathbb{R}^d$. The errors $\{e_{i,\iota}\}$ are assumed to be zero-mean, temporally independent, and spatially correlated, and independent from observations and treatments. The function $f_{\theta_\iota}(\cdot)$ captures spatial spillover effects. This model allows a unit's outcome to depend on the treatments of several units while maintaining conditional independence from treatments in the other units, a common assumption in spatial analysis literature (see e.g., Aronow and Samii, 2017; Reich et al., 2021). Furthermore, one may include the neighborhood-average observation $\bar{O}_{i,\iota}$ in model (1), which is equivalent to redefining the covariate vector as $\tilde{O}_{i,\iota} = (O_{i,\iota}^\top, \bar{O}_{i,\iota}^\top)^\top$. Because we impose no restrictions on the correlation structure among the observations, this reparametrization does not reduce the generality of the model and leaves all theoretical results unchanged.

Under the global design and the linearity assumption that $f_{\theta_\iota}(\{A_{i,j}\}_{j \in \mathcal{N}_i}) = \theta_\iota \bar{A}_{i,\iota}$, different spatial units receive the same treatment and their outcomes satisfy

$$(2) \quad Y_{i,\iota} = \alpha_\iota + O_{i,\iota}^\top \beta_\iota + \gamma_\iota^g A_i + e_{i,\iota},$$

where $\gamma_\iota^g = \gamma_\iota + \theta_\iota$. It is immediate to see that ATE has the closed-form expression $\tau = \sum_{\iota=1}^R \gamma_\iota^g = \sum_{\iota=1}^R (\gamma_\iota + \theta_\iota)$. Using data collected from the global design, we apply the ordinary least squares (OLS) regression to estimate γ_ι^g based on (2) and plug-in these estimators to estimate the ATE, leading to

$$\hat{\tau}^g = \sum_{\iota=1}^R \hat{\gamma}_\iota^g = \sum_{\iota=1}^R u_3^\top \left\{ \sum_{i=1}^N Z_{i,\iota}^g \left(Z_{i,\iota}^g \right)^\top \right\}^{-1} \left(\sum_{i=1}^N Z_{i,\iota}^g Y_{i,\iota} \right),$$

where $u_3 = (0, 0^T, 1)^\top$ and $Z_{i,\iota}^g = (1, O_{i,\iota}^\top, A_i)^\top$.

In the individual-randomized design, we adjust γ_ι^g by splitting it into $\gamma_\iota + \theta_\iota$ and apply OLS to estimate these parameters from the model. The estimators obtained are denoted as $\hat{\gamma}_\iota^i$ and $\hat{\theta}_\iota^i$. For the cluster-randomized design, we define \mathcal{C}_j^0 as the set of ‘‘interior’’ units and $\partial\mathcal{C}_j$ as the ‘‘boundary’’ units of \mathcal{C}_j . A unit is in \mathcal{C}_j^0 if its interference neighbors are all within the same cluster \mathcal{C}_j . In contrast, a unit in boundary $\partial\mathcal{C}_j$ has at least one interference neighbor outside of \mathcal{C}_j . For these boundary units, we estimate the regression coefficients $\hat{\gamma}_\iota^c$ and $\hat{\theta}_\iota^c$ using OLS, similar to the individual-randomized approach. However, in the interior units, γ_ι and θ_ι cannot be identified from each other, leading us to estimate their combined effect using OLS, similar to the global design. This procedure yields the following plug-in estimators:

$$\hat{\tau}^i = \sum_{\iota=1}^R u_{34}^\top \left\{ \sum_{i=1}^N Z_{i,\iota}^i \left(Z_{i,\iota}^i \right)^\top \right\}^{-1} \left(\sum_{i=1}^N Z_{i,\iota}^i Y_{i,\iota} \right),$$

$$\hat{\tau}^c = \sum_{j=1}^m \sum_{\iota \in \mathcal{C}_j} u_\iota^\top \left\{ \sum_{i=1}^N Z_{i,\iota}^c \left(Z_{i,\iota}^c \right)^\top \right\}^{-1} \left(\sum_{i=1}^N Z_{i,\iota}^c Y_{i,\iota} \right),$$

where $u_{34} = (0, 0, 1, 1)^\top$, $Z_{i,\iota}^i = (1, O_{i,\iota}^\top, A_i, \bar{A}_{i,\iota})^\top$, $u_\iota^c = u_{34} \mathbb{I}\{\iota \in \partial\mathcal{C}_j\} + u_3 \mathbb{I}\{\iota \in \mathcal{C}_j^0\}$, and $Z_{i,\iota}^c = (1, O_{i,\iota}^\top, A_i^{(j)}, \bar{A}_{i,\iota})^\top \mathbb{I}\{\iota \in \partial\mathcal{C}_j\} + (1, O_{i,\iota}^\top, A_i^{(j)})^\top \mathbb{I}\{\iota \in \mathcal{C}_j^0\}$.

Semiparametric learning. We now introduce doubly robust (DR) estimators for the ATE, which are widely valued in semiparametric statistics for their resilience to model misspecification (see e.g., [Tsiatis, 2006](#); [Chernozhukov et al., 2018](#)). For the outcome regression, we propose the following model:

$$(3) \quad Y_{i,\iota} = h_\iota(A_{i,\iota}, \bar{A}_{i,\iota}, O_{i,\iota}, \bar{O}_{i,\iota}) + e_{i,\iota},$$

where the error terms $\{e_{i,\iota}\}_{i,\iota}$ s are temporally independent, spatially correlated, and independent of the observed covariates $\{O_{i,\iota}\}$ s and treatments $\{A_{i,\iota}\}$ s.

Model (3) extends our analytical framework in two ways: firstly, it does not constrain the form of h_ι , thereby permitting the use of nonparametric regression or machine learning techniques for estimation; secondly, it incorporates the influence of $\bar{O}_{i,\iota}$, calculated as the average of $O_{i,\iota}$ across interference neighbors. With CA, the ATE can be expressed as

$$\tau = \sum_{\iota=1}^R \mathbb{E}[h_\iota(1, 1, O_{i,\iota}, \bar{O}_{i,\iota}) - h_\iota(0, 0, O_{i,\iota}, \bar{O}_{i,\iota})].$$

To construct the DR estimator, we define the following estimating function for $1 \leq \iota \leq R$, $1 \leq i \leq N$ and $a \in \{0, 1\}$,

$$(4) \quad \nu_{DR}(a, \iota, i, h_\iota, \pi_\iota) = \frac{\mathbb{I}(A_{i,\iota} = a, \bar{A}_{i,\iota} = a)}{\pi_\iota(a|\{O_{i,j}\}_j)} [Y_{i,\iota} - h_\iota(a, a, O_{i,\iota}, \bar{O}_{i,\iota})] + h_\iota(a, a, O_{i,\iota}, \bar{O}_{i,\iota}),$$

Algorithm 1 Doubly robust estimation in nondynamic settings

Require: Data $\{(O_{i,\ell}, A_{i,\ell}, Y_{i,\ell})\}_{i,\ell}$ collected from a given design.

1: Split the data into K ($K \geq 2$) non-overlapped subsets, each with an equal size. Let \mathcal{I}_k denote the indices of the k th data subset.

2: For $k = 1, \dots, K$, compute the estimated outcome regression functions $\{\widehat{h}_\ell^{(k)}\}_\ell$ based on the training data $\{(O_{i,\ell}, A_{i,\ell}, Y_{i,\ell}) : i \notin \mathcal{I}_k, 1 \leq \ell \leq R\}$.

3: If the data is obtained from the individual-randomized design, construct the following ATE estimator

$$\widehat{\tau}_{DR}^i = \frac{1}{N} \sum_{k=1}^K \sum_{i \in \mathcal{I}_k} \sum_{\ell=1}^R \left\{ \nu_{DR}(1, \ell, i, \widehat{h}_\ell^{(k)}, \pi_\ell^i) - \nu_{DR}(0, \ell, i, \widehat{h}_\ell^{(k)}, \pi_\ell^i) \right\}.$$

If the data is obtained from the cluster-randomized design, construct the estimator $\widehat{\tau}_{DR}^c$ by replacing π_ℓ^i with π_ℓ^c in the above equation. Otherwise, compute $\widehat{\tau}_{DR}^g$ by replacing π_ℓ^i with π_ℓ^g .

where $\mathbb{I}(\bullet)$ denotes the indicator function and $\pi_\ell(a|\{O_{i,j}\}_j)$ denotes the propensity score $\mathbb{P}(A_{i,\ell} = a, \bar{A}_{i,\ell} = a|\{O_{i,j}\}_j)$ (explicit expressions are given in Section 1.1 of the Supplementary Material, Yang et al. (2026)). It follows from standard arguments that $\nu_{DR}(a, \ell, i, h_\ell, \pi_\ell)$ is an unbiased estimator for $\mathbb{E}Y_{i,\ell}(\{a\})$ if either the outcome model h_ℓ or the propensity model π_ℓ is correctly specified. The resulting DR estimator of τ is:

$$\frac{1}{N} \sum_{\ell=1}^R \sum_{i=1}^N [\nu_{DR}(1, \ell, i, \widehat{h}_\ell, \pi_\ell) - \nu_{DR}(0, \ell, i, \widehat{h}_\ell, \pi_\ell)],$$

where \widehat{h}_ℓ are estimators of h_ℓ obtained via suitable supervised learning techniques.

To avoid imposing restrictive complexity assumptions on h_ℓ (e.g., metric entropy bounds (Díaz, 2020)), we adopt cross-fitting to compute DR (Chernozhukov et al., 2018). Specifically, the data are partitioned into folds; each fold is used alternately for model training and ATE estimation. The final estimator is obtained by averaging across folds. The complete procedure is detailed in Algorithm 1.

2.3. *Estimation accuracy in the nondynamic setting.* In this subsection, we analyze the MSEs of the ATE estimators under different models and compare different designs.

Parametric estimators. We begin by introducing the following assumption:

ASSUMPTION 2.1. Let $n_\ell^{(j)}$ denote the number of interference neighbors of to the ℓ th unit belonging to the j th cluster. Assume $\omega = \max_{\substack{1 \leq \ell \leq R, \\ 1 \leq j_1, j_2 \leq m}} n_\ell^{(j_1)} / n_\ell^{(j_2)} \mathbb{I}(n_\ell^{(j_2)} > 0) = O(1)$.

Assumption 2.1 requires that, for any unit, the distribution of its interference neighbors across different clusters is not overly uneven, in cases where the unit and its neighbors are not all contained within the same cluster. This condition is mild and typically satisfied in practical settings—particularly when the interference range r is fixed, in which case the number of neighbors per unit is uniformly bounded and the assumption holds automatically. For illustration, Figure 1 displays three example clusterings, where interference neighborhoods are defined based on adjacency. Let $\mathbb{V} = \text{Var}(\vec{e}_i) \in \mathbb{R}^{R \times R}$ where $\vec{e}_i = (e_{i,1}, \dots, e_{i,R})^\top$ denotes the residual vector. We derive the MSEs of ATE estimators under arbitrary treatment probabilities p_ℓ and $p^{(j)}$; full details are presented in Theorem S.1 of the Supplementary Material (Yang et al., 2026). These results imply that MSEs are minimized when $p = p^{(j)} = p_\ell = 0.5$, leading to the following theorem.

THEOREM 2.2. Suppose that CA holds. Set $p = p^{(j)} = p_\ell = 0.5$ for all $1 \leq \ell \leq R$, $1 \leq j \leq m$. Let $r = \max_\ell n_\ell$ and $\nu = \sum_{\ell=1}^R \sum_{\ell'=1}^R \mathbb{V}_{\ell\ell'} / \sum_{\ell=1}^R \mathbb{V}_{\ell\ell}$, it holds that

$$(5) \quad \frac{\text{MSE}(\widehat{\tau}^i)}{\text{MSE}(\widehat{\tau}^g)} \lesssim \frac{(r+1)^2}{\nu},$$

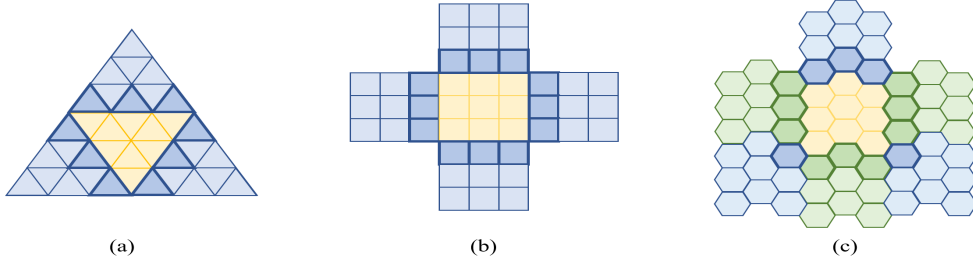


Fig 1: Panels (a), (b), and (c) illustrate examples of clusters with interference neighbors being the adjacent units, each differing in the values of c and r . Specifically, these panels present clusters with $c = 9$ and $r = 3, 4,$ and 6 , respectively. In each panel, the central cluster is emphasized in yellow, with its adjacent units outlined by bold, darker edges. Upon examination, we find that the respective values of ω and the cardinality of $|\mathcal{N}_{\mathcal{C}_j}|$ are $\omega = 1, 3, 3$ and $|\mathcal{N}_{\mathcal{C}_j}| = 9, 12, 14$ for panels (a), (b), and (c), respectively.

where $a_N \lesssim b_N$ means $a_N/b_N = 1 + o(1)$. Further suppose that Assumption 2.1 holds and let $\mathcal{N}_{\mathcal{C}_j} = \cup_{\iota \in \mathcal{C}_j} \mathcal{N}_\iota$. Then it holds that

$$(6) \quad \frac{\text{MSE}(\hat{\tau}^c)}{\text{MSE}(\hat{\tau}^g)} = O\left(\frac{\sum_{j=1}^m \sum_{\iota, \iota' \in \mathcal{C}_j \cup \mathcal{N}_{\mathcal{C}_j}} \mathbb{V}_{\iota \iota'}}{\sum_{\iota, \iota'=1}^R \mathbb{V}_{\iota \iota'}}\right).$$

Theorem 2.2 characterizes how the interference structure influences the MSE of the ATE estimator in spatially randomized experiments. In particular, equation (5) reveals that the relative efficiency of the individual-randomized design over the global design depends on two critical parameters: the interference range r and the correlation strength ν . When the total covariance across units $\sum_{\iota=1}^R \sum_{\iota'=1}^R \mathbb{V}_{\iota \iota'}$ scales linearly with R , that is, $\nu \geq \epsilon R$ for some constant $\epsilon > 0$, we obtain the bound $\text{MSE}(\hat{\tau}^i)/\text{MSE}(\hat{\tau}^g) = O(r^2/R)$. Therefore, when the interference range is small, the individual-randomized design is more capable of generating efficient causal estimator than the global design.

In light of equation (6), the presence of interference effects influences the MSE $\text{MSE}(\hat{\tau}^c)$ not only by the covariance matrices of residuals within the same cluster (e.g., $\{e_\iota\}_{\iota \in \mathcal{C}_j}$), but also by those within interference neighbors of the cluster (e.g., $\{e_\iota\}_{\iota \in \mathcal{N}_{\mathcal{C}_j}}$). To build intuition for the upper bound in (6), we consider two illustrative cases:

Case 1 (Reduction to individual randomization). Suppose that $m = R$ and $\mathcal{C}_j = \{j\}$ for each $1 \leq j \leq m$, so that each unit forms its own cluster. In this case, the cluster-randomized design reduces to the individual-randomized design (Su and Ding, 2021). Equation (6) then simplifies to $O((r+1)^2/\nu)$, which matches the bound in equation (5).

Case 2 (General clustering with bounded correlation). Assume the pairwise covariances $\mathbb{V}_{\iota \iota'}$ are uniformly bounded above, and that the total covariance across all unit pairs, scaled by R^2 , is bounded away from zero. Then, applying the Cauchy-Schwarz inequality, the numerator in (6) satisfies: $\sum_{j=1}^m |\mathcal{C}_j \cup \mathcal{N}_{\mathcal{C}_j}|^2 \leq 2 \sum_{j=1}^m |\mathcal{C}_j|^2 + 2 \sum_{j=1}^m |\mathcal{N}_{\mathcal{C}_j}|^2$. Let $c = \max_j |\mathcal{C}_j|$ denote the largest cluster size. The first term on the right-hand side is then $O(c/R)$. Additionally, if $|\mathcal{N}_{\mathcal{C}_j}|$ is of the order of magnitude $O(r + |\mathcal{C}_j|)$, as is the case in the examples shown in Figure 1, then the second term becomes $O(c/R + mr^2/R^2)$. As such, (6) is bounded by $O(c/R + mr^2/R^2)$.

This result highlights a trade-off: larger clusters increase intra-cluster correlation effects, while broader interference neighborhoods (large r) amplify cross-cluster spillovers. Both inflate the MSE, limiting the efficiency of the cluster-randomized design relative to the global baseline.

Semiparametric estimators. We now examine the theoretical properties of the DR estimator for the ATE. Let $\sigma_O^2 = \text{Var} \left\{ \sum_{\iota=1}^R [h_\iota(1, 1, O_{i,\iota}, \bar{O}_{i,\iota}) - h_\iota(0, 0, O_{i,\iota}, \bar{O}_{i,\iota})] \right\}$ denote the variance component that is independent of the experimental design, and let $\delta_{N,\iota}^2 = \max_{a \in \{0,1\}} \mathbb{E} |\widehat{h}_\iota^{(k)}(a, a, O_{i,\iota}, \bar{O}_{i,\iota}) - h_\iota(a, a, O_{i,\iota}, \bar{O}_{i,\iota})|^2$ represent the estimation error of the regression model for unit ι . In cluster-randomized experiments, the arrangement of a unit and its interference neighbors, whether they fall within the same cluster or are spread across multiple clusters, plays a critical role in determining the DR estimator's performance. To capture this, we partition the residual covariance matrix according to the cluster distribution of interference neighborhoods. Let $r_c(\iota)$ denote the number of clusters that include unit ι and its interference neighbors, and define $r_c = \max_\iota r_c(\iota)$. Let $\mathcal{R} = \{1, 2, \dots, R\}$ as the index set of all units, and $\mathcal{R}_1 = \{\iota : r_c(\iota) > 2\}$. We define

$$\nu_1 = \frac{\sum_\iota \sum_{\iota'} \mathbb{V}_{\iota\iota'}}{\sum_{\iota \in \mathcal{R}_1} \mathbb{V}_{\iota\iota}}, \quad \nu_2 = \frac{\sum_\iota \sum_{\iota'} \mathbb{V}_{\iota\iota'}}{\sum_{\iota \in \mathcal{R}/\mathcal{R}_1} \mathbb{V}_{\iota\iota}},$$

which reflect the relative covariance concentration across these two partitions.

To control the complexity introduced by interference in semiparametric models, we impose the following assumption.

ASSUMPTION 2.3. For each ι , $\sum_{\iota'} \mathbb{I}(1 \leq m_{\iota\iota'} \leq 2) = O(cr)$, where $m_{\iota\iota'} = \sum_{k=1}^m \mathbb{I}((\mathcal{N}_\iota \cup \{\iota\}) \cap \mathcal{C}_k \neq \emptyset, (\mathcal{N}_{\iota'} \cup \{\iota'\}) \cap \mathcal{C}_k \neq \emptyset)$

This condition limits the interference effect across clusters, ensuring that the number of unit pairs with shared or overlapping interference neighborhoods remains manageable. Assumption 2.3 is typically satisfied when interference is distance-based, as in the settings illustrated in Figure 1. We are now ready to present the main result for the DR estimator.

THEOREM 2.4. Suppose that conditions in Theorem 2.2 hold, and $R^{-1} \sum_{\iota=1}^R \delta_{N,\iota}^2 \rightarrow 0$. Then as $N \rightarrow \infty$,

$$(7) \quad \frac{\text{MSE}(\widehat{\tau}_{DR}^i) - N^{-1} \sigma_O^2}{\text{MSE}(\widehat{\tau}_{DR}^g) - N^{-1} \sigma_O^2} \lesssim \frac{(r+1)^2 2^r}{\nu}.$$

When assumption 2.3 holds, we have

$$(8) \quad \frac{\text{MSE}(\widehat{\tau}_{DR}^c) - N^{-1} \sigma_O^2}{\text{MSE}(\widehat{\tau}_{DR}^g) - N^{-1} \sigma_O^2} = O \left(\frac{cr \cdot r_c \cdot 2^{r_c}}{\nu_1} + \frac{cr}{\nu_2} \right).$$

Theorem 2.4 establishes the relative efficiency of DR estimators under individual- and cluster-randomized designs, accounting for spatial interference and estimation error in the outcome regression. Equation (7) shows that, under a fixed interference range r , the MSE ratio between the individual-randomized and global designs is of order $O(R^{-1})$, indicating a potential efficiency gain. However, the ratio grows exponentially with r , due to the importance sampling (IS) $\mathbb{I}(\bar{A}_{i,\iota} = a) / \mathbb{P}(\bar{A}_{i,\iota} = a)$ in (4). We refer such an exponential dependence upon r as the curse of spatial interference. It represents a key limitation of the individual-randomization design under substantial interference. In practice, this issue can be partially mitigated by truncating the denominator $\mathbb{P}(\bar{A}_{i,\iota} = a)$ from below, which avoids excessively large IS weights.

A more principled solution is to adopt cluster-level randomization, whose benefits are reflected in the bound provided by equation (8). The strength of cluster randomization lies in the fact that all units within a cluster receive the same treatment, reducing the variability in

neighborhood treatment assignments. When Assumption 2.3 holds, the curse of interference is largely confined to cluster boundaries, which represent a small fraction of the spatial domain compared to the cluster interiors. The two terms on the right-hand side of equation (8) correspond to contributions from two groups of units. The first term captures the MSE of $\text{MSE}(\hat{\tau}_{DR}^c)$ aggregated over units in \mathcal{R}_1 , to $\text{MSE}(\hat{\tau}_{DR}^g)$. The second term reflects the ratio of $\text{MSE}(\hat{\tau}_{DR}^c)$, aggregated across all units in $\mathcal{R}/\mathcal{R}_1$, whose interference neighborhoods span more than two clusters. Its growth is controlled by the cluster-level interference range r_c , and the exponential term 2^{r_c} reflects the complexity of modeling these interactions. The second term accounts for the remaining units in $\mathcal{R}/\mathcal{R}_1$, which are subject to less interference across cluster boundaries. Importantly, this term lacks the exponential factor 2^{r_c} , and is therefore generally much smaller. These two ratios can be considerably smaller than the RHS of (7) for two reasons: (i) r_c , the cluster-level interference range, can be much smaller than r . For example, in the scenarios illustrated in Figure 1, while r is 3, 4, and 6, r_c is only 3, 3, and 4, respectively. (ii) The second term in (8) is further diminished by a scaling factor $R^{-1} \sum_{j=1}^m |\partial\mathcal{N}_{C_j}|$, which reflects the relatively small proportion of boundary units compared to interior units.

REMARK 1. *The relative efficiency between individual- and cluster-randomized designs can also be inferred from Theorem 2.2. The MSE bounds indicate that the individual-randomized design attains smaller MSE when the number of neighboring units r is small, whereas the cluster-randomized design becomes more efficient as r and the within-cluster dependence increase. Specifically, under the parametric model (1) and the conditions of Theorem 2.2, we have*

$$\frac{\text{MSE}(\hat{\tau}^i)}{\text{MSE}(\hat{\tau}^c)} = O\left(\frac{\sum_{u=1}^R \mathbb{V}_u}{\sum_{j=1}^m \sum_{u,u' \in \mathcal{C}_j \cup \mathcal{N}_{C_j}} \mathbb{V}_{u'}}\right).$$

When all elements of $\mathbb{V}_{u'}$ are positive and uniformly bounded above and below, this simplifies to

$$\frac{\text{MSE}(\hat{\tau}^i)}{\text{MSE}(\hat{\tau}^c)} = O\left(\frac{r^2}{c + mr^2/R}\right).$$

Since $m \approx R/c$, it follows that $\hat{\tau}^c$ converges faster when $r \gg c$, whereas $\hat{\tau}^i$ is more efficient when $r \ll c$. In the semiparametric case (Theorem 2.4), noting that $1/\nu_1 + 1/\nu_2 = 1/\nu$, we similarly find that the cluster-randomized design is preferable when $r \gg \log c$, while the individual-randomized design is favored when $r \ll \log c$.

REMARK 2. *Another implication of Theorem 2.2 concerns the optimal choice of cluster size in cluster-randomized designs. Under the conditions of Theorem 2.2, when all elements of $\mathbb{V}_{u'}$ are positive and uniformly bounded above and below, noting that $m \approx R/c$ and minimizing the leading term $c + mr^2/R$, we obtain that the optimal cluster size satisfies $c^* = r$. The interference range r thus serves as a key design parameter, as it characterizes the spatial extent of spillover effects and directly determines the appropriate granularity of randomization. In the semiparametric case of Theorem 2.4, when the within-cluster interference radius r_c remains constant, smaller clusters generally yield lower MSE. However, if c becomes too small such that the cluster radius is much smaller than the interference radius (i.e., $c \ll r$), then the assumption of constant r_c no longer holds, and cross-cluster interference increases. This transition suggests that the optimal cluster size scales proportionally with the interference range, that is, $c^* \asymp r$, where \asymp denotes asymptotic equivalence up to a constant factor.*

2.4. *Testing power in the nondynamic setting.* In this section, we describe how to construct Wald-type test statistics for ATE inference and analyze how different experimental designs influence the power of these tests. We begin with the asymptotic normality of the estimators.

THEOREM 2.5. *When the conditions of Theorems 2.2 and 2.4 hold,*

(2.5.1) *there exist constants v^g , v^i and v^c such that as $N \rightarrow \infty$,*

$$\sqrt{N(v^g)^{-1}}(\hat{\tau}^g - \tau), \sqrt{N(v^i)^{-1}}(\hat{\tau}^i - \tau), \sqrt{N(v^c)^{-1}}(\hat{\tau}^c - \tau)$$

are asymptotically standard Gaussian distributed.

(2.5.2) *there exist constants v_{DR}^g , v_{DR}^i and v_{DR}^c such that as $N \rightarrow \infty$,*

$$\sqrt{N(v_{DR}^g)^{-1}}(\hat{\tau}_{DR}^g - \tau), \sqrt{N(v_{DR}^i)^{-1}}(\hat{\tau}_{DR}^i - \tau), \sqrt{N(v_{DR}^c)^{-1}}(\hat{\tau}_{DR}^c - \tau)$$

are asymptotically standard Gaussian distributed.

Conclusion (2.5.1) follows from the classical asymptotic normality of the OLS estimator. Conclusion (2.5.2) extends the doubly robust asymptotic theory in Chernozhukov et al. (2018) to accommodate spatial interference. We now turn to the estimation of the variance terms. For the parametric estimators, standard sandwich estimators can be constructed, denoted by $\widehat{\text{Var}}(\hat{\tau}^g)$, $\widehat{\text{Var}}(\hat{\tau}^i)$ and $\widehat{\text{Var}}(\hat{\tau}^c)$. For the doubly robust estimators, the residual covariance $\mathbb{V}_{u'}$ can be estimated in a similar fashion. Let $\Delta_{i\ell} = [\hat{h}_\ell(1, 1, O_{i,\ell}, \bar{O}_{i,\ell}) - \hat{h}_\ell(0, 0, O_{i,\ell}, \bar{O}_{i,\ell})]$ and define $\bar{\Delta}_{i\ell} = N^{-1} \sum_{i=1}^N \Delta_{i\ell}$. The design-invariant variance component σ_O^2 can be estimated via $\hat{\sigma}_O^2 = N^{-1} \sum_{i=1}^N \{\Delta_{i\ell} - \bar{\Delta}_{i\ell}\}^2$. Alternatively, bootstrap procedures can be employed to obtain consistent variance estimates.

To analyze the testing performance under different designs, we consider the following local alternative hypothesis:

$$(9) \quad H_0 : \tau = 0 \text{ V.S. } H_1 : \tau = h/\sqrt{N}.$$

Given a consistent ATE estimator $\hat{\tau}$ and a consistent variance estimate $\widehat{\text{Var}}(\hat{\tau})$, the standard Wald statistic is defined as $\hat{T} = \hat{\tau}/\sqrt{\widehat{\text{Var}}(\hat{\tau})}$. Let c_δ be the $(1 - \delta)$ th quantile of the standard Gaussian distribution and $\Phi(\cdot)$ be the cumulative density function. The null hypothesis in (9) is rejected when $\hat{T} \geq c_\delta$. Under H_0 , Theorem 2.5 implies that $P(\hat{T} \geq c_\delta) \rightarrow \delta$. Under H_1 , we have the following conclusion.

COROLLARY 2.6. *Suppose that the conditions of Theorems 2.2 and 2.4 hold, $\mathbb{V}_{u'}$ is bounded from above and $\sum_{u'} \mathbb{V}_{u'}/R^2$ is bounded from zero. Under the alternative hypothesis in (9), as $N \rightarrow \infty$,*

(2.6.1) *the sufficient and necessary conditions for test consistency are:*

$$P(\hat{T}^g \geq c_\delta) \rightarrow 1 \iff h \gg c_\delta R,$$

$$P(\hat{T}^i \geq c_\delta) \rightarrow 1 \iff h \gg c_\delta r \sqrt{R},$$

$$P(\hat{T}^c \geq c_\delta) \rightarrow 1 \iff h \gg c_\delta(\sqrt{mr} + \sqrt{cR}).$$

(2.6.2) for the DR estimators, the corresponding conditions are:

$$\begin{aligned} P\left(\widehat{T}_{DR}^g \geq c_\delta\right) &\rightarrow 1 \iff h \gg c_\delta \sqrt{\sigma_O^2 + R^2}, \\ P\left(\widehat{T}_{DR}^i \geq c_\delta\right) &\rightarrow 1 \iff h \gg c_\delta \sqrt{\sigma_O^2 + r2^r R}, \\ P\left(\widehat{T}_{DR}^c \geq c_\delta\right) &\rightarrow 1 \iff h \gg c_\delta \{\sigma_O^2 + |\mathcal{R}_1| c r r_c 2^{r_c} + (R - |\mathcal{R}_1|) c r\}^{1/2}. \end{aligned}$$

These results highlight the comparative efficiency of spatially randomized designs. From (2.6.1), if the interference range r is fixed and the cluster size $c \ll R$, then both $\{r\sqrt{R}, \sqrt{mr} + \sqrt{cR}\}$ grow at a slower rate than R , implying that the spatial designs achieve consistency at smaller signal strengths. Similarly, in (2.6.2), when r and r_c are constants, $c \ll R$ and the number of boundary regions $|\mathcal{R}_1| \ll R$ (e.g., $c = O(R^\delta)$ for some $0 < \delta < 1$), we have $r2^r R \ll R^2$, and $|\mathcal{R}_1| c r r_c 2^{r_c} + (R - |\mathcal{R}_1|) c r \ll R^2$. Under these mild conditions, the power of tests under spatially randomized designs converges to 1 at a much faster rate than under global randomization. This implies that, for a fixed signal strength, spatial randomization offers greater testing efficiency, particularly in large-scale experiments.

3. Dynamic Setting. In this section, we consider a dynamic environment where, within each day, each spatial unit receives a sequence of M (possibly different) treatments over time. The observed data take the form $\{(A_{iut}, O_{iut}, Y_{iut})\}_{1 \leq i \leq N, 1 \leq \iota \leq R, 1 \leq t \leq M}$ where i indexes the i th day, ι indexes the ι th unit, t indexes the t th time interval, corresponding to i.i.d. copies of $\{(A_{iut}, O_{iut}, Y_{iut})\}_{1 \leq \iota \leq R, 1 \leq t \leq M}$. To simplify the presentation, we have chosen not to employ the potential outcome framework to formulate the causal estimand. Readers who are interested in exploring the potential outcomes approach may refer to the works by [Luckett et al. \(2019\)](#) and [Shi et al. \(2023a\)](#). The ATE in this context is given by:

$$\tau = \sum_{t=1}^M \sum_{\iota=1}^R [\mathbb{E}_1(Y_{iut}) - \mathbb{E}_0(Y_{iut})],$$

where \mathbb{E}_1 and \mathbb{E}_0 denote the expectation assuming that treatments are assigned according to two non-dynamic policies: consistently set to 1 and consistently set to 0, respectively. In settings with multiple decision stages, past interventions may influence current variables, thereby indirectly impacting current outcomes. This carryover effects adds complexity to both the estimation process and theoretical analysis.

3.1. *Parametric and semiparametric modeling.* For parametric modeling, in addition to the outcome regression model, we posit another first-order autoregression model for the observational variable, yielding the following models:

$$(10) \quad \begin{aligned} Y_{iut} &= \alpha_{it} + O_{iut}^\top \beta_{it} + \gamma_{it} A_{iut} + \theta_{it} \bar{A}_{iut} + e_{iut}, \\ O_{i,t+1} &= \Lambda_{it} + B_{it} O_{iut} + \Gamma_{it} A_{iut} + \Theta_{it} \bar{A}_{iut} + E_{iut}, \end{aligned}$$

where $\Lambda_{it}, \Gamma_{it}, \Theta_{it} \in \mathbb{R}^d$, $B_{it} \in \mathbb{R}^{d \times d}$, $\bar{A}_{iut} = n_\iota^{-1} \sum_{k \in \mathcal{N}_\iota} A_{ikt}$ and $\{e_{iut}, E_{iut}\}$ are i.i.d. copies of $\{e_{it}, E_{it}\}$ which are independent of $\{O_{iut}\}$ and satisfy $\mathbb{E}(e_{it}) = 0$, $\text{cov}(e_{it}, e_{\iota't'}) = \mathbb{V}_{\iota'\iota}^e \mathbb{I}\{t = t'\}$, $\mathbb{E}(E_{it}) = 0$, and $\text{cov}(E_{it}, E_{\iota't'}) = \mathbb{V}_{\iota'\iota}^E \mathbb{I}\{t = t'\}$, respectively, and for all ι , e_{it} and E_{it} are independent over t . This assumption rules out temporal dependence in the error structure, which, if present, may induce endogeneity and complicate identification and inference. Addressing such time-dependent endogeneity is beyond the scope of this work; see [Luo et al. \(2024\)](#) for possible solutions for such settings.

For the global design where $A_{iut} = A_{it}$ for any ι , model (10) degenerates to

$$(11) \quad \begin{aligned} Y_{iut} &= \alpha_{it} + O_{iut}^\top \beta_{it} + \gamma_{it}^g A_{it} + e_{iut}, \\ O_{i,t+1} &= \Lambda_{it} + B_{it} O_{iut} + \Gamma_{it}^g A_{it} + E_{iut}, \end{aligned}$$

where $\gamma_{it}^g = \gamma_{it} + \theta_{it}$ and $\Gamma_{it}^g = \Gamma_{it} + \Theta_{it}$. We now turn our focus to estimating τ within the framework of models (10) and (11). We first present the subsequent proposition.

PROPOSITION 3.1. *Define c_{it}^\top as $\sum_{k=t+1}^M \beta_{ik}^\top \left(\prod_{j=t+1}^{k-1} B_{ij} \right)$. Under models (10) and (11), the ATE can be expressed as*

$$\tau = \sum_{\iota=1}^R \sum_{t=1}^M (\gamma_{it}^g + c_{it}^\top \Gamma_{it}^g) = \sum_{\iota=1}^R \sum_{t=1}^M \{\gamma_{it} + \theta_{it} + c_{it}^\top (\Gamma_{it} + \Theta_{it})\}.$$

Proposition 3.1 shows that the ATE consists of two components: the initial terms $\gamma_{it}^c = \gamma_{it} + \theta_{it}$ represents the direct effect of A_t on the immediate outcomes, whereas the last term $c_{it}^\top \Gamma_{it}^g = c_{it}^\top (\Gamma_{it} + \Theta_{it})$ corresponds to the indirect effect, measuring the delayed treatment effect on future outcomes through O_{t+1} . Analogous to the nondynamic setting, one can derive OLS estimators for the regression coefficients and plug-in them into the expression in Proposition 3.1 to obtain the estimated τ . To save space, we do not detail the estimating procedure again.

For semiparametric modeling, we adopt the double reinforcement learning (DRL) framework of Kallus and Uehara (2020), assuming a time-varying Markov decision process (MDP) (Puterman, 2014). The Markov property implies that, conditional on the current state and action $(\mathbf{O}_t, \mathbf{A}_t)$, future observations and immediate outcomes are independent of the past. Here, for $a \in \{0, 1\}$, let the target policy π_a denote the deterministic policy that sets all actions to a . The Q-function is defined as $Q_{it}^a(\mathbf{O}_t, \mathbf{A}_t) = \sum_{k=t}^M \mathbb{E}_a(Y_{ik} | \mathbf{O}_t, \mathbf{A}_t)$, where the current action A_{it} is fixed by conditioning, and the expectation is taken over future transitions $k > t$ under the policy π_a . We also define the density ratio between target and behavior policies as

$$\mu_{it}^a(\mathbf{O}_t, \mathbf{A}_t) = \frac{\prod_{\iota} \mathbb{I}(A_{it} = a) p_a(\mathbf{O}_t)}{\mathbb{P}(\cap_{\iota} \{A_{it} = a\} | \mathbf{O}_t) p_b(\mathbf{O}_t)}, \quad a \in \{0, 1\},$$

where p_a and p_b denote the density functions of \mathbf{O}_t under the target policy and the behavior policy respectively. In dynamic settings, the state distribution evolves according to the transition law $p(\mathbf{O}_{t+1} | \mathbf{O}_t, \mathbf{A}_t)$. Because past actions affect subsequent states, the marginal state distribution under the target policy may differ from under the behavior policy. Accordingly, the ratio μ_{it}^a accounts for both the difference in action probabilities and the policy-dependent state distributions. The resulting DRL estimator is given by

$$\begin{aligned} \tilde{\tau}_{DRL} = & \frac{1}{N} \sum_{a=0}^1 (-1)^{a+1} \sum_{i=1}^N \sum_{\iota=1}^R \left\{ Q_{i1}^a(\mathbf{O}_{i1}, a) \right. \\ & \left. + \sum_{t=1}^M \mu_{it}^a(\mathbf{O}_{it}, \mathbf{A}_{it}) [Y_{iit} + Q_{it+1}^a(\mathbf{O}_{it+1}, a) - Q_{it}^a(\mathbf{O}_{it}, \mathbf{A}_{it})] \right\}. \end{aligned}$$

The main challenge lies in estimating Q_{it}^a and μ_{it}^a , which depend on the high-dimensional input $(\mathbf{O}_t, \mathbf{A}_t) \in \mathbb{R}^{2Rd}$. Moreover, the variance of $\tilde{\tau}_{DRL}$ may grow exponentially with the input dimension, akin to the ‘‘curse of horizon’’ (Liu et al., 2018). To address this, we adopt the mean-field approximation (Yang et al., 2018; Shi et al., 2023b). For each unit ι , define $X_{it} = (O_{it}, A_{it}, m_{\iota}(\mathbf{O}_t, \mathbf{A}_t))$, where $m_{\iota}(\cdot)$ summarizes local neighborhood information, e.g., averages over interference neighbors. We assume: (i) Q_{it}^a and $\mathbb{E}(Y_{it} | \mathbf{O}_t, \mathbf{A}_t)$ depend only on X_{it} ; (ii) $O_{\iota, t+1}$ and $m_{\iota}(\cdot)$ are conditionally independent of past history given X_{it} . These assumptions are testable using modern Markov or conditional independence tests (Chen and Hong, 2012; Zhou et al., 2023b; Zhang et al., 2011). Under these assumptions, it can be shown that $\tilde{\tau}_{DRL}$ is unbiased to the following (Shi et al., 2023b)

$$\hat{\tau}_{DRL} = \frac{1}{N} \sum_{a=0}^1 (-1)^{a+1} \sum_{i=1}^N \sum_{\iota=1}^R \left\{ Q_{i1}^a(X_{i1}^a) + \sum_{t=1}^M \mu_{it}^a(X_{iit}) [Y_{iit} + Q_{it+1}^a(X_{it+1}^a) - Q_{it}^a(X_{iit})] \right\},$$

where $X_{it}^a = (O_{it}, a, m_\iota(\mathbf{O}_t, a))$. Unlike $\tilde{\tau}_{DRL}$, this estimator relies on low-dimensional inputs and avoids exponential variance growth.

It remains to estimate Q_{it}^a and μ_{it}^a to compute $\hat{\tau}_{DRL}$. To estimate Q_{it}^a , we employ backward induction (Murphy, 2003). Specifically, we begin by estimating $Q_{\iota M}^a(X_{\iota M}) = \mathbb{E}(Y_{\iota M} | X_{\iota M})$. Let $\hat{Q}_{\iota M}^a$ denote the resulting estimator. We next recursively estimate $Q_{it}^a(X_{it}) = \mathbb{E}[Y_{it} + \hat{Q}_{it+1}^a(X_{it+1}^a) | X_{it}]$ for $t = M - 1, M - 2, \dots, 1$ and construct these estimators $\{\hat{Q}_{it}^a\}_t$ via nonparametric regression (e.g., kernel smoothing or splines). To estimate μ_{it}^a , we employ minimax learning (Liu et al., 2018; Uehara, Huang and Jiang, 2020). To save space, we relegate the detailed implementation to Section 1.2 of the Supplementary Material (Yang et al., 2026). Finally, similar to the contextual bandit setting, we employ data-splitting and cross-fitting to construct $\hat{\tau}_{DRL}$. We summarize our procedure in Algorithm 2.

Algorithm 2 Doubly reinforcement learning in dynamic settings

Require: Data $\{(O_{i,\iota,t}, A_{i,\iota,t}, Y_{i,\iota,t})\}_{i,\iota,t}$ collected from a given design.

1: Split all data trajectories into K ($K \geq 2$) non-overlapped subsets, each with an equal size. Let \mathcal{D}_k denote the indices of days that belong to the k th data subset where $k = 1, 2, \dots, K$.

2: For $k = 1, \dots, K$, $a \in \{0, 1\}$, $1 \leq t \leq M$ and $1 \leq \iota \leq R$, compute $\hat{\mu}_{it}^{a,(k)}$ and $\hat{Q}_{it}^{a,(k)}$ based on data trajectories $\{1, \dots, n\} \setminus \mathcal{D}_k$.

3: Set

$$\hat{\tau}_{DRL} = \frac{1}{N} \sum_{a=0}^1 (-1)^{a+1} \sum_{k=1}^K \sum_{i \in \mathcal{D}_k} \sum_{\iota=1}^R \left\{ \hat{Q}_{i1}^{a,(k)}(X_{i1}^a) + \sum_{t=1}^M \hat{\mu}_{it}^{a,(k)}(X_{it}) [Y_{i\iota t} + \hat{Q}_{i\iota t+1}^{a,(k)}(X_{i\iota t+1}^a) - \hat{Q}_{i\iota t}^{a,(k)}(X_{i\iota t})] \right\}.$$

3.2. *Estimation accuracy in the dynamic setting.* Since treatments are sequentially assigned over time, we incorporate the spatial designs with the following temporal designs: (i) The *constant design* that sets each unit the same treatment at each day, i.e. $A_{i1} = A_{i2} = \dots = A_{iM}$ for each $1 \leq \iota \leq R$. (ii) The *independent design* in which all $\{A_{i\iota t}\}_t$ are independent for any t and ι . (iii) The *switchback design* that switches the treatments for each unit back and forth (the initial treatment on each day is randomly generated), i.e. $A_{i1} = 1 - A_{i2} = A_{i3} = 1 - A_{i4} = \dots$ for any ι .

We first compare different designs under the parametric model assumption (10). Let $\hat{\tau}_{OLS}^g, \hat{\tau}_{OLS}^i$ and $\hat{\tau}_{OLS}^c$ be the estimators under the global, individual- and cluster-randomized designs, respectively. Define the composite noise term $u_{it} = c_{it}^\top E_{iit} + e_{iit}$, where c_{it} is defined the same as in Proposition 3.1 and represents the delayed indirect effects. Let $\mathbb{V}_{\iota' t}^u$ denote $\text{cov}(u_{i\iota t}, u_{i\iota' t})$, and define the heterogeneity index

$$\eta = \sum_{t=1}^M \sum_{\iota=1}^R \sum_{\iota'=1}^R \mathbb{V}_{\iota' t}^u / \sum_{t=1}^M \sum_{\iota=1}^R \mathbb{V}_{\iota t}^u.$$

THEOREM 3.2. *Under any of the three temporal designs (constant, independent, or switchback), it holds that*

$$\frac{N \cdot \text{MSE}(\hat{\tau}_{OLS}^i) - \sigma_{OLS}^2}{N \cdot \text{MSE}(\hat{\tau}_{OLS}^g) - \sigma_{OLS}^2} \lesssim \frac{(r+1)^2}{\eta}.$$

Furthermore, if Assumption 2.1 holds, we have

$$\frac{N \cdot \text{MSE}(\hat{\tau}_{OLS}^c) - \sigma_{OLS}^2}{N \cdot \text{MSE}(\hat{\tau}_{OLS}^g) - \sigma_{OLS}^2} = O\left(\frac{\sum_{j=1}^m \sum_{\iota, \iota' \in \mathcal{C}_j \cup \mathcal{N}_{\mathcal{C}_j}} \mathbb{V}_{\iota' t}^u}{\sum_{\iota, \iota'} \mathbb{V}_{\iota' t}^u}\right).$$

The results in Theorem 3.2 are consistent with those in Theorem 2.2. The explicit expression of σ_{OLS}^2 is presented in Section 3.6 of the Supplementary Material (Yang et al., 2026). We remark that if the indirect effects are weak, e.g., $\Gamma_{it} = O(R^{-1/2})$ for any i and t , or if for any t , $\{E_{it}\}_i$ are spatially correlated only within interference neighborhoods, we obtain

$$\frac{\sigma_{OLS}^2}{N \cdot \text{MSE}(\hat{\tau}_{OLS}^g) - \sigma_{OLS}^2} = O\left(\frac{r^2}{R}\right).$$

Moreover, under standard regularity conditions, such as boundedness of \mathbb{V}_{it}^u and nonvanishing aggregate variance, we also have

$$\frac{\text{MSE}(\hat{\tau}_{OLS}^i)}{\text{MSE}(\hat{\tau}_{OLS}^g)} = O\left(\frac{r^2}{R}\right) \quad \text{and} \quad \frac{\text{MSE}(\hat{\tau}_{OLS}^c)}{\text{MSE}(\hat{\tau}_{OLS}^g)} = O\left(\frac{c}{R} + \frac{mr^2}{R^2}\right).$$

To compare the semiparametric estimators, we assume:

ASSUMPTION 3.3. For $1 \leq i \leq R$, $1 \leq t \leq M$ and $a = 0, 1$, the reward Y_{it} and the density ratio μ_{it}^a are bounded; the sample splitting estimators $\hat{\mu}_{it}^{a,(k)}$ and $\hat{Q}_{it}^{a,(k)}$ are finite.

Denote the DRL ATE estimators under the global, individual- and cluster-randomized designs as $\hat{\tau}_{DRL}^g$, $\hat{\tau}_{DRL}^i$ and $\hat{\tau}_{DRL}^c$, respectively. Let $u_{it} = r_{it} - \mathbb{E}(r_{it}|O_{it}, A_{it}, m_i(\mathbf{O}_{it}, \mathbf{A}_{it}))$ be the noise term and $C_{i_1, i_2, t} = \text{cov}\{u_{i_1, t}, u_{i_2, t}\}$ be its covariance. Then we have the following result in parallel to Theorem 2.4.

THEOREM 3.4. Suppose that $m_i(\mathbf{O}_t, \mathbf{A}_t) = (n_i^{-1} \sum_{k \in \mathcal{N}_i} O_{kt}, n_i^{-1} \sum_{k \in \mathcal{N}_i} A_{kt})$. Under the condition that r is finite, and all elements in $\{C_{i_1, i_2, t} : 1 \leq i_1, i_2 \leq R, 1 \leq t \leq M\}$ are nonnegative with a substantial proportion distinctly greater than zero, the following result holds as $N \rightarrow \infty$:

$$\frac{N \cdot \text{MSE}(\hat{\tau}_{DRL}^i) - \sigma_{DRL}^2}{N \cdot \text{MSE}(\hat{\tau}_{DRL}^g) - \sigma_{DRL}^2} = O\left(\frac{r^2}{R^2}\right).$$

where $\sigma_{DRL}^2 = \text{Var}\{\sum_{i=1}^R (Q_{i1}(\mathbf{1}_{RM}) - Q_{i1}(\mathbf{0}_{RM}))\}$. Furthermore, if Assumption 2.3 holds, then

$$\frac{N \cdot \text{MSE}(\hat{\tau}_{DRL}^c) - \sigma_{DRL}^2}{N \cdot \text{MSE}(\hat{\tau}_{DRL}^g) - \sigma_{DRL}^2} = O\left(\frac{cr \cdot r_c \cdot 2^{r_c}}{\nu_1} + \frac{cr}{\nu_2}\right).$$

Similar to the nondynamic case, we establish the asymptotic normality of the ATE estimators in the dynamic framework when the number of stages M is finite. See Theorem S.4 in the Supplementary Material (Yang et al., 2026) for details. Results regarding hypothesis testing follows directly from this result. For brevity, we omit these results in the main text.

4. Numerical Experiments. In this section, we conduct numerical experiments to validate the theoretical insights derived from Sections 2-3. We focus on scenarios where interference neighbors are defined as adjacent neighboring units. This setup is particularly relevant in contexts such as ride-sharing markets, where the impact of a policy in one area can influence neighboring units through driver distribution, under the assumption that drivers can only move to adjacent areas within a given time frame.

4.1. *Simulation of the nondynamic setting.* We evaluate performance in a single-stage setting with $R = 36, 81,$ and 144 units, arranged as regular polygons with $r = 3, 4,$ and 6 sides, respectively. Clusters are of fixed size $|\mathcal{C}_j| = 9$, yielding $m = 4, 9,$ and 16 clusters as depicted in Figure 2.

In the parametric model (1), covariates $O_{i,\iota}$ are i.i.d following $N(4, 1)$. Each unit ι has normalized coordinates $(\iota_x, \iota_y) \in (0, 1)^2$. We set $\alpha_\iota = 8 + 2\{f_1(\iota_x) + g_1(\iota_y)\}$ and $\beta_\iota = f_2(\iota_x) + g_2(\iota_y)$, where f_k, g_k are Fourier series of the form $a_0 + \sum_{k=1}^K (a_k \cos k\pi x + b_k \sin k\pi x)$ with $K = 3$ and coefficients drawn from $U(0, 1)$. Treatment effects follow $\gamma_\iota = s_y \cdot \alpha_\iota / \sum_\iota \alpha_\iota$ and $\theta_\iota = 0.6s_y \cdot \beta_\iota / \sum_\iota \beta_\iota$, where s_y corresponds to $s\%$ improvement over the control outcome. We vary $s \in \{0, 0.25, 0.5, \dots, 2\}$.

For the semi-parametric model, the underlying model is

$$Y_{t,\iota} = 5 + 3(O_{i,\iota} + \bar{O}_{i\iota}) \sin \left\{ \frac{\pi}{8}(\iota_x + \iota_y) + s \cdot A_{t,\iota} + 0.5 \cdot s \cdot \bar{A}_{i,\iota} \right\} + 0.5e_{i\iota},$$

where $O_{i,\iota} \sim N(4, 1)$ truncated to $(3, 5)$ and s varies from 0 to 0.015 in increments of 0.0025. The noise terms $e_{i\iota}$ are mean-zero with $\text{cov}(e_\iota, e_{\iota'}) = \rho^{d_{\iota\iota'}}$, where $d_{\iota\iota'} = \sqrt{(x_\iota - x_{\iota'})^2 + (y_\iota - y_{\iota'})^2} / 2$ is the scaled spatial distance. We vary $\rho \in \{0.3, 0.6, 0.9\}$, with $N = 30$ samples per configuration and 500 replications. Additional noise structures are discussed in Section 4 of the Supplementary Material (Yang et al., 2026).

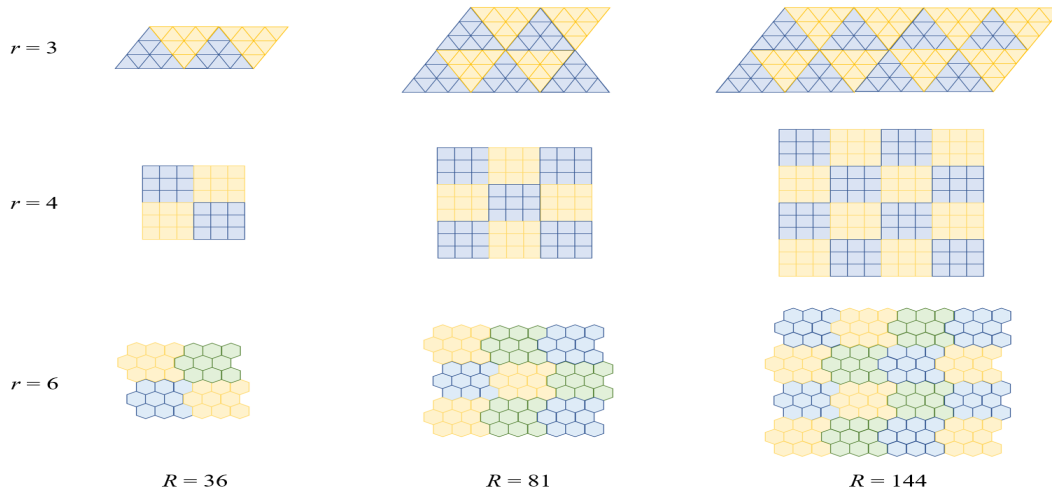


Fig 2: Simulation layout showcasing unital patterns. Each row illustrates configurations with a maximum number of neighbors per unit, specifically $r = 3, 4,$ and 6 . The columns indicate varying total counts of units, with $R = 36, 81,$ and 144 . Within each panel, distinct colors denote separate clusters, leading to varying cluster counts of $m = 4, 9,$ and 16 .

Table 1 presents empirical MSE ratios, where r_1 and r_2 denote the ratios of the individual- and cluster-randomized designs to the global design, respectively, for both parametric ($\hat{\tau}$) and semiparametric ($\hat{\tau}_{DR}$) estimators. The results demonstrate that spatially randomized designs consistently reduce MSEs compared to the global design, especially as the noise correlation ρ and/or the number of units R increases. The individual-randomized design is generally more efficient than the cluster-randomized one when the number of neighbors r is small (e.g., $r = 3, 4$), while the cluster-randomized design becomes preferable for larger r (e.g., $r = 6$).

Inference results based on these settings show that both designs control type I error near the nominal 0.05 level when $s = 0$. As the treatment effect increases, spatially randomized designs exhibit superior power under both models. Figure 3 further confirms that power improves with larger R or smaller r , aligning with our theoretical findings.

TABLE 1
 Nondynamic setting: MSE ratio comparisons under parametric and semiparametric models across different (r, ρ) settings.

Model	R	Ratio	(3,0.9)	(3,0.6)	(3,0.3)	(4,0.9)	(4,0.6)	(4,0.3)	(6,0.9)	(6,0.6)	(6,0.3)
Parametric	36	r_1	0.360	0.379	0.421	0.544	0.582	0.661	0.982	1.042	1.171
		r_2	0.522	0.554	0.593	0.641	0.690	0.759	0.674	0.724	0.795
	64	r_1	0.214	0.200	0.195	0.316	0.306	0.319	0.577	0.557	0.577
		r_2	0.370	0.361	0.361	0.457	0.460	0.485	0.493	0.497	0.524
	144	r_1	0.117	0.130	0.155	0.187	0.219	0.259	0.357	0.413	0.485
		r_2	0.185	0.196	0.212	0.246	0.272	0.090	0.271	0.298	0.339
Semiparametric	36	r_1	0.186	0.202	0.239	0.352	0.382	0.459	0.286	0.316	0.385
		r_2	0.306	0.331	0.378	0.438	0.484	0.576	0.363	0.399	0.466
	64	r_1	0.103	0.106	0.118	0.165	0.175	0.204	0.138	0.153	0.185
		r_2	0.163	0.173	0.192	0.199	0.214	0.244	0.190	0.203	0.231
	144	r_1	0.062	0.068	0.079	0.096	0.107	0.127	0.080	0.090	0.107
		r_2	0.089	0.095	0.107	0.102	0.111	0.129	0.102	0.111	0.129

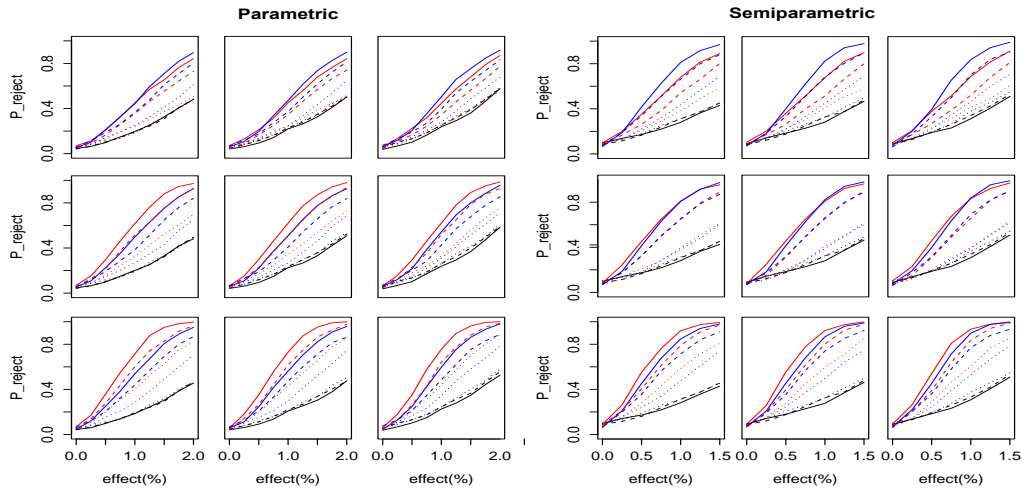


Fig 3: Rejection probabilities in the parametric and semiparametric models under a non-dynamic setup. The horizontal axis represents the treatment's relative improvement. The red lines depict the individual-randomized design, the blue lines represent the cluster-randomized design, and the black lines denote the global randomized design. The line styles—solid, dashed, and dotted—correspond to the number of units $R = 144, 81,$ and $36,$ respectively. The figure is organized into three rows and three columns of panels, representing different values of r (6, 4, and 3) and ρ (0.9, 0.6, and 0.3), respectively.

4.2. *Simulation of the dynamic setting.* In the dynamic setting, we use the same unit configurations as in the static case. For the parametric model, initial covariates $O_{i1} \sim N(4, 1)$ are i.i.d., and for $t \geq 1$, covariates O_{it} follow model (10). Coefficients $\alpha_{it}, \beta_{it}, \Lambda_{it}$ vary over time via functions $h_1(t), h_2(t), h_3(t)$:

$$\alpha_{it} = h_1(t)\{8 + 2(f_1(\iota_x) + g_1(\iota_y))\}, \beta_{it} = h_2(t)\{f_2(\iota_x) + g_2(\iota_y)\}, \Lambda_{it} = h_3(t)\{f_3(\iota_x) + g_3(\iota_y)\},$$

where each $h_i(t)$ is of the form $a_0 + \sum_{k=1}^K (a_k \cos(k\pi t) + b_k \sin(k\pi t))$ with $a_0, a_k, b_k \sim U(0, 1), K = 3$. Similarly, define $B_{it}^0 = h_4(t)f_4(\iota_x) + g_4(\iota_y)$, then normalize: $B_{it} = 0.5(B_{it}^0 - \min B^0) / (\max B^0 - \min B^0) + 0.3$. Noises e_{it} are i.i.d. as in the static case, while E_{it} are i.i.d. copies of $0.1e_{\iota}$. Effect coefficients are scaled by signal strengths s_x and s_y : $\gamma_{it} = s_y \cdot \alpha_{it} / \sum_{\iota} \sum_t \alpha_{it}$, $\theta_{it} = 0.6s_y \cdot \beta_{it} / \sum_{\iota} \sum_t \beta_{it}$, $\Gamma_{it} = s_x \cdot \Lambda_{it} / \sum_{\iota} \sum_t \Lambda_{it}$ and $\Theta_{it} = 0.6s_x \cdot B_{it} / \sum_{\iota} \sum_t B_{it}$, with $s_x = s \cdot \sum_{\iota, t} \mathbb{E}(O_{it} | A_{jk} = 0)$ and $s_y = s^0 \cdot \sum_{\iota, t} \mathbb{E}(Y_{it} | A_{jk} = 0)$ representing $s\%$ improvements.

For the nonparametric model, the underlying model is

$$Y_{it} = 5 + 2 \cdot O_{it} \sin \left\{ \frac{\pi}{8} (\iota_x + \iota_y + t/M) + s \cdot (A_{it} + \bar{A}_{it}) \right\} + 0.5e_{it},$$

where $(O_{i1t}, \dots, O_{iRt})^\top$ are i.i.d. multivariate Gaussians with mean $(2 + A_{i1t}, \dots, 2 + A_{iRt})^\top$ and covariance $\Sigma(\rho)$ having unit variances and off-diagonals $v_\iota v_\iota^\top$, where $v_\iota \sim \text{Unif}(0.75, 1)$ for about a ρ .

We simulate $r = 3, 4, 6$ (number of neighbors), $\rho = 0.3, 0.6, 0.9$ (noise correlation), $R = 36, 81, 144$ (units), and $M = 12$ time steps, each with $N = 30$ observations. Table 2 summarizes empirical MSE ratios across settings, showing consistent trends with the static case and supporting Theorems 3.2 and 3.4. Figure 4 presents power curves for various s levels. In all scenarios, spatially randomized designs outperform the global design, highlighting their effectiveness in dynamic environments fraction of indices and 0 otherwise, yielding a low-rank correlation structure controlled by ρ .

TABLE 2

Dynamic setting: MSE ratio comparisons under parametric and semiparametric models across different (r, ρ) settings.

Model	R	Metric	(3, 0.9)	(3, 0.6)	(3, 0.3)	(4, 0.9)	(4, 0.6)	(4, 0.3)	(6, 0.9)
Parametric	36	r_1	0.315	0.350	0.410	0.551	0.603	0.694	1.067
		r_2	0.495	0.513	0.543	0.693	0.738	0.808	0.726
	64	r_1	0.249	0.252	0.267	0.434	0.444	0.475	0.709
		r_2	0.291	0.302	0.326	0.388	0.416	0.471	0.428
	144	r_1	0.116	0.126	0.142	0.196	0.211	0.238	0.366
		r_2	0.103	0.114	0.138	0.158	0.179	0.224	0.165
Semiparametric	36	r_1	0.139	0.135	0.152	0.181	0.156	0.181	0.193
		r_2	0.641	0.713	0.809	0.666	0.757	0.871	0.705
	64	r_1	0.073	0.070	0.075	0.084	0.082	0.085	0.089
		r_2	0.425	0.474	0.587	0.355	0.409	0.503	0.352
	144	r_1	0.033	0.032	0.033	0.037	0.035	0.038	0.043
		r_2	0.224	0.250	0.296	0.218	0.245	0.291	0.221

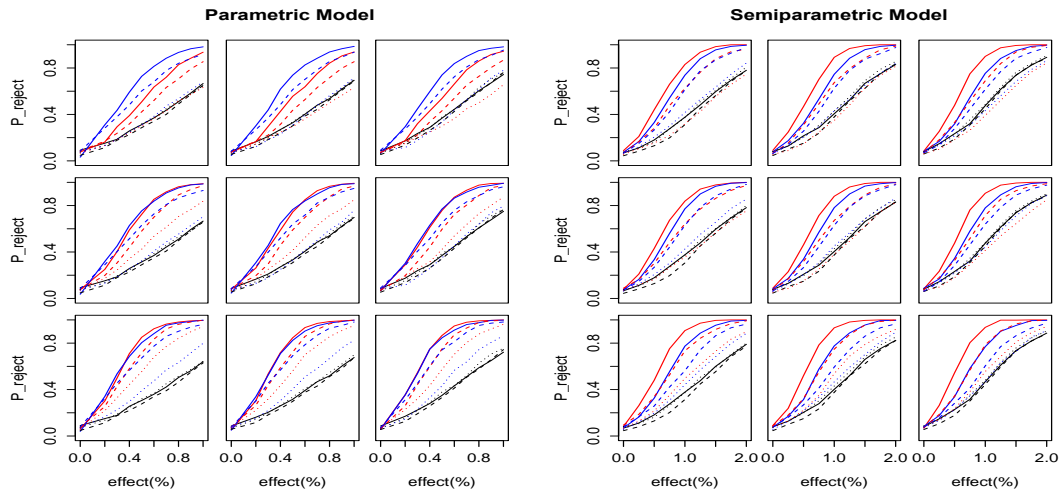


Fig 4: Rejection probability in the parametric and nonparametric models with $M = 12$ of dynamic setting under different relative improvements of the new policy. The red, blue and black lines represent to the individual-, cluster- and global-randomized designs, with $R = 144, 81, 36$ plotted in solid, 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.

4.3. *Real data based simulation.* In this section, we evaluate the performance of our proposed experimental framework by simulating a real online ride-hailing service, using the same simulator in Zhou et al. (2021). This simulator accurately models the dynamic interplay between demand and supply, mirroring an actual ride-hailing platform. It generates demand distributions from historical data, while initial driver supply distributions are based on early-day data, subsequently adjusted through the simulator’s dynamic transitions and order dispatch algorithms. The simulation’s key metrics, such as driver earnings, response rates, and idle times, deviate from real-world figures by less than 2%.

For our analysis, we selected the largest square unit of the city, dividing it into 64 equally sized, non-overlapping squares ($R = 64$). Over $N = 20$ days, segmented into $M = 24$ hourly intervals, we investigated the effects of policies simulating ATE improvements of 0%, 1.0%, 2%, and 5%. We applied a parametric regression model for the multi-stage scenario to estimate and analyze the ATE, comparing the classical global design with an individual randomized design that incorporates temporal randomization. The outcomes of these designs are denoted by $\hat{\tau}_{OLS}^i$ and $\hat{\tau}_{OLS}^g$, respectively.

Using 300 Monte Carlo simulations to closely estimate the true ATE, we compiled the test powers and MSE ratios $\text{MSE}(\hat{\tau}_{OLS}^i)/\text{MSE}(\hat{\tau}_{OLS}^g)$ in Table 3. Our findings indicate that both the classical global and spatially randomized designs maintain proper type I error rates. Notably, the spatially randomized design consistently exhibited lower MSE and higher test power compared to the classical global approach, irrespective of interference effects. This evidence strongly supports the superior efficiency of the spatially randomized design in real-world applications.

TABLE 3
Test powers and MSE ratios $\text{MSE}(\hat{\tau}_{OLS}^i)/\text{MSE}(\hat{\tau}_{OLS}^g)$ under different effect sizes and interference settings.

Effect Size	ATE-test (No Interference)		ATE-test (With Interference)		MSE Ratio	
	Benchmark	Proposed	Benchmark	Proposed	No Interf.	Interf.
0%	0.07	0.06	0.07	0.05	0.010	0.129
1%	0.16	1.00	0.16	0.32	0.013	0.134
2%	0.33	1.00	0.33	0.89	0.016	0.128
5%	0.87	1.00	0.87	1.00	0.037	0.158

4.4. *Real data example.* To illustrate our approach, we analyze data from a dynamic A/B experiment conducted on a ride-sourcing platform to evaluate a customer-subsidy policy. The study area was partitioned into $R = 17$ regions, and treatments were assigned via an individual-randomized switchback design over $N = 24$ days (2020/02/19–2020/03/13), with 30-minute decision intervals ($M = 48$ per day). For interference, we defined each region’s adjacent regions as its neighbors, consistent with common assumptions in spatial transportation settings. This choice is supported empirically: over 90% of ride requests were served by drivers within the same region, implying that interactions are predominantly local. We take the number of requests as the state variable and drivers’ total income as the outcome. While the spatial dimension here is modest relative to the rich temporal resolution, the setting provides a concrete and realistic demonstration of how our framework can be implemented and assessed in dynamic, operational environments, including cases where the number of spatial units is limited.

Applying our parametric dynamic model, we detect a statistically significant treatment effect ($p = 1.0158 \times 10^{-5}$), suggesting that the customer subsidy increased overall driver income. Figure 5 displays the observed and fitted income trajectories for two economically

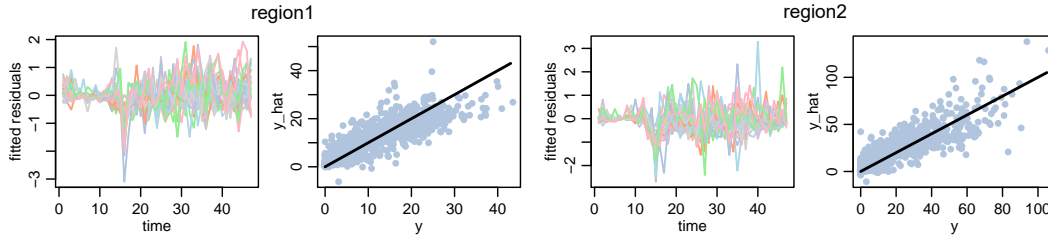


Fig 5: Plots of the corresponding relative residuals, as well as the fitted drivers' income against the observed values.

dominant regions that contribute the most to citywide revenue, along with their relative residuals. Here, the relative residual is defined as the fitted residual divided by the standard deviation of the corresponding region's income. In these high-activity regions, the fitted trajectories closely follow the observed daily dynamics, and the residuals exhibit no systematic patterns, indicating a good fit where the signal is strongest. By contrast, several low-income regions with sparse ride requests show larger residual variability, plausibly reflecting weaker signal strength and higher relative noise. Because these regions contribute little to the aggregate outcome, we highlight the dominant regions for clarity and space economy. Overall, this case study illustrates how the proposed framework can be used to evaluate dynamic spatial policies in large-scale online experiments.

5. Concluding Remarks. In this paper, we aim at enhancing the efficiency of inference across various scenarios by assigning treatment randomly across units. A pivotal discovery is that randomization effectively diminishes the spatial correlation of noise. Under standard conditions, we observe that the MSE ratio of ATE estimators, derived from spatially randomized designs compared to classical global designs, is inversely related to the number of units or clusters, applicable in both static and dynamic frameworks.

Moving forward, there are several important topics worthy of further investigation. Firstly, examining experimental designs with temporal random effects is an interesting issue, which may cause endogeneity bias in the state autoregression model. One possible solution is to assume that random effects satisfy certain covariance structures such as declining correlation as temporal distance increases. Secondly, investigating more complex interference structures and/or more complex treatment variables (Ao, Calonico and Lee, 2021; Dong, Lee and Gou, 2023) would be an interesting problem.

6. Significance Statement. Modern online platforms, such as ride-sharing and digital marketplaces, rely heavily on A/B testing to evaluate policies, yet these systems exhibit strong spatial and temporal interactions that standard experimental designs often ignore. This paper develops a unified framework for designing and analyzing experiments under such interference. We show that spatially randomized designs can substantially improve estimation accuracy and testing power, and we provide practical guidance on when to use individual or cluster randomization and how to choose cluster size. Our results bridge theory and practice by linking experimental design to real-world policy evaluation, and are demonstrated using data from a large-scale ride-sourcing platform.

Acknowledgments. Hongtu Zhu is the corresponding author. E-mail: htzhu@email.unc.edu.

Funding. This research is supported by the National Natural Science Foundation of China (No. 12292981, 72301276, 12571309, 62588101, 11931001), National Key R&D Program of China grants (No. 2022YFA1003800), the New Cornerstone Science Foundation through Xplorer Prize, the LMAM and the Fundamental Research Funds for the Central Universities (LMEQF) and an EPSRC grant EP/W014971/1.

SUPPLEMENTARY MATERIAL

Supplementary Material for “Spatially Randomized Designs Can Enhance Policy Evaluation” (DOI:[10.1214/\[providedbytypesetter\]-AOAS2173SUPPA](https://doi.org/10.1214/[providedbytypesetter]-AOAS2173SUPPA);pdf). The additional implementation details, simulation results and supplementary theoretical results along with the proofs unshown in the main text.

SpatiallyRandomization-code (DOI:[10.1214/\[providedbytypesetter\]-AOAS2173SUPPB](https://doi.org/10.1214/[providedbytypesetter]-AOAS2173SUPPB);zip). Code files to generate the numerical results in the main text. The data can be found on the GitHub page at <https://github.com/anneyang0060/SpatiallyRandomization>. The real data used in our study is proprietary and cannot be shared publicly. However, we have provided a synthetic dataset that can yield similar results for broader accessibility and understanding.

REFERENCES

- AO, W., CALONICO, S. and LEE, Y.-Y. (2021). Multivalued treatments and decomposition analysis: An application to the WIA program. *Journal of Business & Economic Statistics* **39** 358–371.
- ARONOW, P. M. and SAMII, C. (2017). Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics* **11** 1912–1947.
- BASSE, G., DING, P., FELLER, A. and TOULIS, P. (2024). Randomization tests for peer effects in group formation experiments. *Econometrica* **92** 567–590.
- CALLAWAY, B. and LI, T. (2023). Policy evaluation during a pandemic. *Journal of Econometrics* **236** 105454.
- CHEN, B. and HONG, Y. (2012). Testing for the Markov property in time series. *Econometric Theory* **28** 130–178.
- CHERNOZHUKOV, V., CHETVERIKOV, D., DEMIRER, M., DUFLO, E., HANSEN, C., NEWEY, W. and ROBINS, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal* **21**.
- DAI, R., WANG, J., ZHOU, F., LUO, S., QIN, Z., SHI, C. and ZHU, H. (2024). Causal deepsets for off-policy evaluation under spatial or spatio-temporal interferences. *arXiv preprint arXiv:2407.17910*.
- DÍAZ, I. (2020). Machine learning in the estimation of causal effects: targeted minimum loss-based estimation and double/debiased machine learning. *Biostatistics* **21** 353–358.
- DING, P. (2024). *A first course in causal inference*. CRC Press.
- DONG, Y., LEE, Y.-Y. and GOU, M. (2023). Regression discontinuity designs with a continuous treatment. *Journal of the American Statistical Association* **118** 208–221.
- FORASTIERE, L., AIROLDI, E. M. and MEALLI, F. (2021). Identification and estimation of treatment and interference effects in observational studies on networks. *Journal of the American Statistical Association* **116** 901–918.
- HANNA, J., NIEKUM, S. and STONE, P. (2019). Importance sampling policy evaluation with an estimated behavior policy. In *International Conference on Machine Learning* 2605–2613. PMLR.
- HU, Y. and HU, F. (2012). Asymptotic Properties of Covariate-adaptive Randomization. *The Annals of Statistics* **40** 1794–1815.
- HU, Y., LI, S. and WAGER, S. (2022). Average direct and indirect causal effects under interference. *Biometrika*.
- HUBER, M. and STEINMAYR, A. (2021). A framework for separating individual-level treatment effects from spillover effects. *Journal of Business & Economic Statistics* **39** 422–436.
- HUDGENS, M. G. and HALLORAN, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association* **103** 832–842.
- IMAI, K., TINGLEY, D. and YAMAMOTO, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society Series A: Statistics in Society* **176** 5–51.
- IMBENS, G. W. and RUBIN, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- JAGADEESAN, R., PILLAI, N. S. and VOLFOVSKY, A. (2020). Designs for estimating the treatment effect in networks with interference. *The Annals of Statistics* **48** 679–712.
- JIA, S., KALLUS, N. and YU, C. L. (2023). Clustered switchback experiments: Near-optimal rates under spatiotemporal interference. *arXiv preprint arXiv:2312.15574*.
- JIANG, N. and LI, L. (2016). Doubly robust off-policy value evaluation for reinforcement learning. In *International Conference on Machine Learning* 652–661. PMLR.
- 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.

- KALLUS, N. and UEHARA, M. (2022). Efficiently breaking the curse of horizon in off-policy evaluation with double reinforcement learning. *Operations Research* **70** 3282–3302.
- KE, J., YANG, H., ZHENG, H., CHEN, X., JIA, Y., GONG, P. and YE, J. (2018). Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services. *IEEE Transactions on Intelligent Transportation Systems* **20** 4160–4173.
- KONG, X., YUAN, M. and ZHENG, W. (2021). Approximate and exact designs for total effects. *The Annals of Statistics* **49** 1594–1625.
- LARSEN, A., YANG, S., REICH, B. J. and RAPPOLD, A. G. (2022). A spatial causal analysis of wildland fire-contributed PM_{2.5} using numerical model output. *The Annals of Applied Statistics* **16** 2714–2731.
- LEUNG, M. P. (2022). Rate-optimal cluster-randomized designs for spatial interference. *The Annals of Statistics* **50** 3064–3087.
- LI, T., SHI, C., WANG, J., ZHOU, F. et al. (2023). Optimal Treatment Allocation for Efficient Policy Evaluation in Sequential Decision Making. *Advances in Neural Information Processing Systems* **36**.
- LIU, Y. and HU, F. (2022). Balancing unobserved covariates with covariate-adaptive randomized experiments. *Journal of the American Statistical Association* **117** 875–886.
- LIU, L., HUDGENS, M. G. and BECKER-DREPS, S. (2016). On inverse probability-weighted estimators in the presence of interference. *Biometrika* **103** 829–842.
- 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**.
- LU, X., LIU, T., LIU, H. and DING, P. (2023). Design-based theory for cluster rerandomization. *Biometrika* **110** 467–483.
- LUCKETT, D. J., LABER, E. B., KAHKOSKA, A. R., MAAHS, D. M., MAYER-DAVIS, E. and KOSOROK, M. R. (2019). Estimating dynamic treatment regimes in mobile health using v-learning. *Journal of the American Statistical Association* **115** 692–706.
- LUEDTKE, A. R. and VAN DER LAAN, M. J. (2016). Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Annals of statistics* **44** 713.
- LUO, S., YANG, Y., SHI, C., YAO, F., YE, J. and ZHU, H. (2024). Policy Evaluation for Temporal and/or Spatial Dependent Experiments. *Journal of the Royal Statistical Society, Series B* **86** 623–649.
- MA, W., LI, P., ZHANG, L.-X. and HU, F. (2024). A new and unified family of covariate adaptive randomization procedures and their properties. *Journal of the American Statistical Association* **119** 151–162.
- MURPHY, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **65** 331–355.
- PEREZ-HEYDRICH, C., HUDGENS, M. G., HALLORAN, M. E., CLEMENS, J. D., ALI, M. and EMCH, M. E. (2014). Assessing effects of Cholera vaccination in the presence of interference. *Biometrics* **70** 731–741.
- PUELZ, D., BASSE, G., FELLER, A. and TOULIS, P. (2022). A graph-theoretic approach to randomization tests of causal effects under general interference. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **84** 174–204.
- PUTERMAN, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.
- REICH, B. J., YANG, S., GUAN, Y., GIFFIN, A. B., MILLER, M. J. and RAPPOLD, A. (2021). A review of spatial causal inference methods for environmental and epidemiological applications. *International Statistical Review* **89** 605–634.
- SELTMAN, H. J. (2012). *Experimental design and analysis*.
- SHI, C., WAN, R., CHERNOZHUKOV, V. and SONG, R. (2021). Deeply-debiased off-policy interval estimation. In *International Conference on Machine Learning* 9580–9591. PMLR.
- SHI, C., ZHANG, S., LU, W. and SONG, R. (2022). Statistical inference of the value function for reinforcement learning in infinite-horizon settings. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **84** 765–793.
- SHI, C., WANG, X., LUO, S., ZHU, H., YE, J. and SONG, R. (2023a). Dynamic causal effects evaluation in a/b testing with a reinforcement learning framework. *Journal of the American Statistical Association* **118** 2059–2071.
- SHI, C., WAN, R., SONG, G., LUO, S., ZHU, H. and SONG, R. (2023b). A multiagent reinforcement learning framework for off-policy evaluation in two-sided markets. *The Annals of Applied Statistics* **17** 2701–2722.
- SOBEL, M. E. (2006). What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *Journal of the American Statistical Association* **101** 1398–1407.
- SU, F. and DING, P. (2021). Model-assisted analyses of cluster-randomized experiments. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **83** 994–1015.
- SUN, K., KONG, L., ZHU, H. and SHI, C. (2024). Arma-design: Optimal treatment allocation strategies for a/b testing in partially observable time series experiments. *arXiv preprint arXiv:2408.05342*.

- TANG, X., QIN, Z., ZHANG, F., WANG, Z., XU, Z., MA, Y., ZHU, H. and YE, J. (2019). A deep value-network based approach for multi-driver order dispatching. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* 1780–1790.
- TCHETGEN, E. J., FULCHER, I. R. and SHPITSER, I. (2021). Auto-g-computation of causal effects on a network. *Journal of the American Statistical Association* **116** 833–844.
- THOMAS, P. and BRUNSKILL, E. (2016). Data-efficient off-policy policy evaluation for reinforcement learning. In *International Conference on Machine Learning* 2139–2148. PMLR.
- TSIATIS, A. A. (2006). *Semiparametric theory and missing data*. Springer.
- 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.
- UEHARA, M., SHI, C. and KALLUS, N. (2026). A Review of Off-Policy Evaluation in Reinforcement Learning. *Statistical Science* in press.
- WANG, J., QI, Z. and WONG, R. K. (2023). Projected state-action balancing weights for offline reinforcement learning. *The Annals of Statistics* **51** 1639–1665.
- WANG, J., WEN, Q., ZHANG, Y., YAN, X. and SHI, C. (2025). A two-armed bandit framework for a/b testing. *arXiv preprint arXiv:2507.18118*.
- WEN, Q., SHI, C., YANG, Y., TANG, N. and ZHU, H. (2025). Unraveling the Interplay between Carryover Effects and Reward Autocorrelations in Switchback Experiments. In *International Conference on Machine Learning* 66481–66519. PMLR.
- WU, X., WEN, Q., ZHANG, Y., ZHU, H., LI, T. and SHI, C. (2026). Designing Time Series Experiments in A/B Testing with Transformer Reinforcement Learning. *arXiv preprint arXiv:2602.01853*.
- XU, Z., LI, Z., GUAN, Q., ZHANG, D., LI, Q., NAN, J., LIU, C., BIAN, W. and YE, J. (2018). Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 905–913.
- YANG, Y., LUO, R., LI, M., ZHOU, M., ZHANG, W. and WANG, J. (2018). Mean field multi-agent reinforcement learning. In *International conference on machine learning* 5571–5580. PMLR.
- YANG, Y., SHI, C., YAO, F., WANG, S. and ZHU, H. (2026). Supplement to “Spatially Randomized Designs Can Enhance Policy Evaluation”. [https://doi.org/DOI:10.1214/\[providedbytypesetter\]](https://doi.org/DOI:10.1214/[providedbytypesetter])
- YUAN, Y., ALTENBURGER, K. and KOOTI, F. (2021). Causal network motifs: Identifying heterogeneous spillover effects in a/b tests. In *Proceedings of the Web Conference 2021* 3359–3370.
- ZHANG, W., YANG, Y. and YAO, F. (2026). High-Dimensional Detection of Spatial Interference Effects. *Journal of Business & Economic Statistics* **0** 1–12. <https://doi.org/10.1080/07350015.2025.2581341>
- ZHANG, K., PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2011). Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence* 804–813.
- ZHOU, F., LUO, S., QIE, X., YE, J. and ZHU, H. (2021). Graph-based equilibrium metrics for dynamic supply–demand systems with applications to ride-sourcing platforms. *Journal of the American Statistical Association* **116** 1688–1699.
- ZHOU, W., LI, Y., ZHU, R. and QU, A. (2023a). Distributional shift-aware off-policy interval estimation: A unified error quantification framework. *arXiv preprint arXiv:2309.13278*.
- ZHOU, Y., SHI, C., LI, L. and YAO, Q. (2023b). Testing for the Markov property in time series via deep conditional generative learning. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **85** 1204–1222.
- ZHOU, Y., LIU, Y., LI, P. and HU, F. (2024). Cluster-adaptive network a/b testing: From randomization to estimation. *Journal of Machine Learning Research* **25** 1–48.
- ZHOU, H., HANNA, J. P., ZHU, J., YANG, Y. and SHI, C. (2025). Demystifying the paradox of importance sampling with an estimated history-dependent behavior policy in off-policy evaluation. *arXiv preprint arXiv:2505.22492*.
- ZHU, J., LI, J., ZHOU, H., LIN, Y., LIN, Z. and SHI, C. (2025). Balancing Interference and Correlation in Spatial Experimental Designs: A Causal Graph Cut Approach. *arXiv preprint arXiv:2505.20130*.
- ZIGLER, C. M., DOMINICI, F. and WANG, Y. (2012). Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics* **13** 289–302.
- ZIGLER, C. M. and PAPADOGEORGOU, G. (2021). Bipartite causal inference with interference. *Statistical Science: a Review Journal of the Institute of Mathematical Statistics* **36** 109.