Causal Inference with Noncompliance and Unknown Interference

Tadao Hoshino* and Takahide Yanagi†

August, 2021

Abstract

In this paper, we investigate a treatment effect model in which individuals interact in a social network and they may not comply with the assigned treatments. We introduce a new concept of exposure mapping, which summarizes spillover effects into a fixed dimensional statistic of instrumental variables, and we call this mapping the instrumental exposure mapping (IEM). We investigate identification conditions for the intention-to-treat effect and the average causal effect for compliers, while explicitly considering the possibility of misspecification of IEM. Based on our identification results, we develop nonparametric estimation procedures for the treatment parameters. Their asymptotic properties, including consistency and asymptotic normality, are investigated using an approximate neighborhood interference framework by Leung (2021). For an empirical illustration of our proposed method, we revisit Paluck *et al.*'s (2016) experimental data on the anti-conflict intervention school program.

Keywords: exposure mapping; instrumental variables; local average treatment effect; network interference; spillover effects.

JEL Classification: C14, C31, C51.

*School of Political Science and Economics, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo 169-8050, Japan. Email: thoshino@waseda.jp.

†Graduate School of Economics, Kyoto University, Yoshida Honmachi, Sakyo, Kyoto, 606-8501, Japan. Email: yanagi@econ.kyoto-u.ac.jp.

1 Introduction

In recent years, we have witnessed increasing importance in evaluating causal effects under cross-unit interference in many fields. When individuals interact with each other, using the conventional potential outcome framework of Rubin (1980) based on the *stable unit treatment value assumption* (SUTVA) is inappropriate. To address the potential interference, there has been a rapidly growing number of studies that attempt to mitigate SUTVA by replacing it with some weaker restrictions on the interference structure. For comprehensive reviews of the literature on this issue, see VanderWeele and An (2013), Halloran and Hudgens (2016), and Aronow *et al.* (2021).

A common approach to dealing with interference is to assume the existence of a low-dimensional *exposure mapping* which serves as a sufficient statistic of spillover effects in that others' treatments affect one's outcomes only through this function (e.g., Hong and Raudenbush, 2006; Hudgens and Halloran, 2008; Manski, 2013; Aronow and Samii, 2017; Li *et al.*, 2019; Egami, 2021; Forastiere *et al.*, 2021; among others). Some frequently-used forms of exposure mapping include, for example, simply extracting treatment vectors of disjoint groups, or calculating the proportion of treated neighbors. The exposure mapping is a useful tool for summarizing potentially complicated spillover effects, but there is an inherent difficulty in how to choose the "right" functional form. Thus, some recent studies attempt to uncover under what conditions one can estimate meaningful treatment parameters even under *unknown interference*, in which the exposure mapping is misspecified or is not explicitly specified (Chin, 2018; Leung, 2021; Sävje, 2021; Sävje *et al.*, 2021).

Including the aforementioned studies, much of the research on causal inference with interference assumes the availability of experimental data where the individuals fully comply with their assigned treatments. However, this should be restrictive in many applications (e.g., Miguel and Kremer, 2004; Dupas, 2014; Zelizer, 2019; among many others). As a real example, consider the experiment on social norms and behavioral patterns of adolescents conducted by Paluck *et al.* (2016). They randomly selected students to participate in the anti-conflict intervention program where the participants were encouraged to take on leadership roles to reduce conflict behaviors in school. The authors were interested in assessing the effectiveness of the intervention against one's own behavior, as well as whether the participants influence their peers through their social network. Unfortunately, a proportion of the selected students did not join the intervention program, which led the authors to resort to an intention-to-treat (ITT) analysis that counts non-compliers as directly treated students.

Although the coexistence of spillovers and noncompliance should be prevalent in relevant empirical applications, only a few studies have explicitly tackled this issue. Sobel (2006) shows that the conventional estimators, such as the two-stage least squares estimator, do not admit causal interpretations when ignoring spillover effects. In addition, there are recent studies that develop solutions to this problem based on the instrumental variable (IV) method by extending the local average treatment effect framework of Imbens and Angrist (1994) and Angrist *et al.* (1996) (e.g., Kang and Imbens, 2016; Imai *et al.*, 2021; DiTraglia *et al.*, 2021; Vazquez-Bare, 2021). However, these studies focus only on the situations where spillovers occur within disjoint clusters and, more importantly, they do not address the issue of potential misspecification of the exposure mapping.

Taking these points into account, it should be of primary importance to understand what treatment parameters we can identify (if any) and how to make statistical inferences on them under the possibility of noncompliance and unknown interference structure. This is the objective of this study, that is, to develop a new causal inference method that allows for noncompliance while retaining robustness to the misspecification of exposure mapping. We consider a model in which individuals are connected through a social network and they may self-select their treatment status. To account for the noncompliance issue and network interference, we employ the IV method and introduce a new concept of exposure mapping, which we call *instrumental exposure mapping* (IEM). The IEM is similar to the conventional exposure mapping in that it is a function summarizing the spillover effects into low-dimensional variables, but it differs in that it is a function of IVs.

We first examine the ITT analysis, in which the estimands of interest are the average direct effect (ADE) and average spillover effect (ASE) of the IV on the outcome and the treatment choice. We show that ADE has a clear causal interpretation even under the misspecification of IEM, however, the misspecification gives rise to a difficulty in interpreting ASE. Next, we focus on the identification of the ADE for *compliers* who comply with their assigned treatments, which we call the *local average direct effect* (LADE). Under certain conditions, LADE captures the ADE of the treatment receipt on the outcome for compilers, and thus LADE should be more interpretable and policy-relevant than the simple ITT parameter. The technical difficulty in identifying LADE is that, without imposing any restrictions on the interaction structure, the conventional identification conditions for models without spillovers do not ensure the identification. To address this problem, we extend the restricted interference assumption by Imai *et al.* (2021) to our situation under network interference by restricting the spillover effects caused by non-compliers in a certain way. It is shown that the LADE parameter can be identified by a Wald-type estimand under this restricted interference assumption.

Given our identification results, we propose nonparametric estimation procedures for the ITT effect and LADE. The proposed estimators are easy to implement, although their statistical properties are non-trivial due to network dependence. We provide a set of sufficient conditions for our estimators to be consistent and asymptotically normally distributed by utilizing the approximate neighborhood interference (ANI) framework by Leung (2021). We adopted the ANI framework because (i) ANI is suitable for many empirical situations where spillovers from distant units are sufficiently small but potentially nonzero, and (ii) ANI is conceptually similar to the familiar *near-epoch dependence* condition in networks, making it relatively easy to interpret and verify. Technically, ANI ensures that the data satisfy the ψ-*weak dependence*, and we can employ the limit theorems for ψ-weakly dependent processes (Kojevnikov *et al.*, 2021; Kojevnikov, 2021) to derive the asymptotic properties of our estimators. We also consider statistical inference methods based on a network HAC estimation and network-dependent bootstrapping.

As an empirical illustration, we apply our method to Paluck *et al.*'s (2016) data, the experimental data on the anti-conflict intervention program for adolescents. The results of the ITT analysis show that receiving an invitation to the intervention program has a statistically significant positive effect on the students' own anti-conflict norms and behaviors. While this finding is consistent with previous studies, the LADE estimates suggest even larger treatment effects. Thus, the ITT analysis might have underestimated the direct effects of the intervention program, highlighting the importance of estimating the LADE.

Related literature Our identification results build especially on Imai *et al.* (2021) who consider the identification of average causal effects for compliers, which they call the *complier average direct effect* and *complier average spillover effect*, in two-stage randomized experiments under noncompliance. An important assumption underlying their model is that interference is restricted within disjoint groups (such as classrooms or rural villages), differently from our focus on network interference. More importantly, they require the exposure mapping to be correctly specified by the stratified interference assumption in which spillovers are determined only through the number of the assigned treatments within each cluster.

Another study closely related to ours is Leung (2021). The paper proposes an ANI model in experimental situations with perfect compliance and develops inference methods for average treatment effect parameters while explicitly allowing for the misspecification of exposure mapping. The major distinction between Leung (2021) and ours is that not only the spillover effects of treatments on the outcome but also the spillovers of others' IVs on own treatment choice are considered in our study. The existence of two interference channels significantly complicates the identification analysis and statistical inference.

This paper also relates to prior studies that have extended Fisher's (1935) randomization test to the cases with interference (e.g., Athey *et al.*, 2018; Basse *et al.*, 2019; Li *et al.*, 2019). Similar to our approach, the interference structure is generally left unspecified in these studies, and randomization tests can be implemented without requiring the specification of exposure mapping. However, the main focus of these studies is on testing some hypotheses regarding the spillover effects, while the primary purpose of this paper is to establish procedures for the identification and estimation of the treatment parameters.

Paper organization In Section 2, we present basic model assumptions with some examples. Section 3 provides the identification results. We discuss the estimation and inference procedures in Section 4, where their asymptotic properties are also investigated. Section 5 reports the numerical results including the Monte Carlo experiments and empirical analysis. Section 6 concludes the paper. Appendix A contains the proofs of all technical results in the main text. The other supplementary results are relegated to Appendix B.

2 Model

Consider a finite population of $n \in \mathbb{N}$ units $N_n := \{1, 2, \dots, n\}$, where $\mathbb{N} = \{1, 2, \dots\}$. Suppose that the units form an undirected and unweighted network. The network is represented by the $n \times n$ symmetric adjacency matrix $\boldsymbol{A} = (A_{ij})_{i,j \in N_n}$, where $A_{ij} \in \{0, 1\}$ indicates whether or not i and j are adjacent, i.e., $A_{ij} = 1$ if there is a link between i and j and $A_{ij} = 0$ otherwise. As usual, we assume that there are no self-links so that $A_{ii} = 0$ for all $i \in N_n$. We denote the set of possible adjacency matrices of n units as $\mathcal{A}_n$.

In a later section, we study asymptotic theory under the condition that the network size n grows to infinity. This means that we consider a sequence of networks $\{\boldsymbol{A}_m\}_{m \in \mathbb{N}}$, where $\boldsymbol{A}_m \in \mathcal{A}_m$ and m can be any large number. The observed adjacency matrix $\boldsymbol{A}$ with no subscript is regarded as an n-th element of the sequence, i.e., $\boldsymbol{A} = \boldsymbol{A}_n$. The networks $\boldsymbol{A}_{m_1}$ and $\boldsymbol{A}_{m_2}$ ($m_1 \neq m_2$) can be completely unrelated, but it may be possible that $\boldsymbol{A}_{m_1+m_2}$ is a union of disjoint $\boldsymbol{A}_{m_1}$ and $\boldsymbol{A}_{m_2}$. Further, the distributions of variables including potential outcomes and treatments can be specific to each network in general; that is, they form a triangular array defined along with the network sequence. However, for notational simplicity, we suppress the dependence of variables on the network structure.

Let $Y_i \in \mathbb{R}$ be an observed outcome variable and $D_i \in \{0, 1\}$ an indicator of the treatment receipt for unit $i \in N_n$. In observational studies or randomized experiments with possible noncompliance, individuals may self-select their treatment status and the existing methods under perfect compliance are generally not applicable.

To address this problem, suppose that there is a binary IV, $Z_i \in \{0,1\}$. In an experimental setup, Z_i is typically an indicator of initial treatment recommendation for i. Denote the n-dimensional vector of realized treatments as $\boldsymbol{D} = (D_i)_{i \in N_n}$, and similarly let $\boldsymbol{Z} = (Z_i)_{i \in N_n}$. We write the support of $\boldsymbol{D}$ and that of $\boldsymbol{Z}$ as $\mathcal{D}_n = \{0,1\}^n$ and $\mathcal{Z}_n = \{0,1\}^n$, respectively. For each $\boldsymbol{d} \in \mathcal{D}_n$ and $\boldsymbol{z} \in \mathcal{Z}_n$, we denote $Y_i(\boldsymbol{d}, \boldsymbol{z}) \in \mathbb{R}$ as the potential outcome of unit i when $\boldsymbol{D} = \boldsymbol{d}$ and $\boldsymbol{Z} = \boldsymbol{z}$. Similarly, the potential treatment status given $\boldsymbol{Z} = \boldsymbol{z}$ is written as $D_i(\boldsymbol{z}) \in \{0,1\}$. Let $\boldsymbol{D}(\boldsymbol{z}) = (D_i(\boldsymbol{z}))_{i \in N_n}$ be the n-dimensional vector of potential treatments. By construction, we have $Y_i = Y_i(\boldsymbol{D}, \boldsymbol{Z})$, $D_i = D_i(\boldsymbol{Z})$, and $\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{Z})$. Hence, we can further write $y_i(\boldsymbol{z}) = Y_i(\boldsymbol{D}(\boldsymbol{z}), \boldsymbol{z})$ for some function $y_i : \mathcal{Z}_n \to \mathbb{R}$, and we have

$$Y_i = y_i(\boldsymbol{Z}) = \sum_{\boldsymbol{z} \in \mathcal{Z}_n} \mathbf{1}\{\boldsymbol{Z} = \boldsymbol{z}\} y_i(\boldsymbol{z}).$$

Because we can observe only one realization from $(y_i(\boldsymbol{z}), D_i(\boldsymbol{z}))_{\boldsymbol{z} \in \mathcal{Z}_n}$ for each unit, it is generally impossible to define identifiable causal estimands without introducing some restrictions. To address this issue, we introduce the following function: $T : N_n \times \mathcal{Z}_n \times \mathcal{A}_n \to \mathcal{T}$, where T is a pre-specified function, $\mathcal{T} \subset \mathbb{R}^{\dim(T)}$ is a set which does not depend on i and n, and $\dim(T)$ is a fixed positive integer.[1] For each i, we denote $T_i = T(i, \boldsymbol{Z}, \boldsymbol{A})$. We call the function T the *instrumental exposure mapping* (IEM) and its realization T_i the *instrumental exposure* of unit i.

Denote $\boldsymbol{z}_{-i} = (z_j)_{j \neq i}$ and $\boldsymbol{z} = (z_i, \boldsymbol{z}_{-i})$. We say that the IEM is *correctly specified* if for any $i \in N_n$, $z_i \in \{0,1\}$, $\boldsymbol{z}_{-i}, \boldsymbol{z}'_{-i} \in \{0,1\}^{n-1}$, and $\boldsymbol{A} \in \mathcal{A}_n$,[2]

$$T(i, z_i, \boldsymbol{z}_{-i}, \boldsymbol{A}) = T(i, z_i, \boldsymbol{z}'_{-i}, \boldsymbol{A}) \Longrightarrow D_i(z_i, \boldsymbol{z}_{-i}) = D_i(z_i, \boldsymbol{z}'_{-i}) \text{ and } y_i(z_i, \boldsymbol{z}_{-i}) = y_i(z_i, \boldsymbol{z}'_{-i}). \tag{2.1}$$

Thus, if the IEM is correctly specified, it serves as a fixed dimensional sufficient statistic that summarizes potentially high-dimensional spillover effects. That is, the potential treatment status and the potential outcome of unit i can be fully characterized by i's own IV Z_i and her instrumental exposure T_i, and there exist functions $\widetilde{d}_i : \{0,1\} \times \mathcal{T} \to \{0,1\}$ and $\widetilde{y}_i : \{0,1\} \times \mathcal{T} \to \mathbb{R}$ satisfying

$$\widetilde{d}_i(z_i, T(i, z_i, \boldsymbol{z}_{-i}, \boldsymbol{A})) = D_i(z_i, \boldsymbol{z}_{-i}) \text{ and } \widetilde{y}_i(z_i, T(i, z_i, \boldsymbol{z}_{-i}, \boldsymbol{A})) = y_i(z_i, \boldsymbol{z}_{-i})$$

for any $z_i \in \{0,1\}$ and $\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}$. Then, $\widetilde{y}_i(z,t)$ and $\widetilde{d}_i(z,t)$ represent the potential outcome and the potential treatment status, respectively, given $Z_i = z$ and $T_i = t$. In this way, a properly specified IEM alleviates the complexity of handling general spillover effects and greatly simplifies the estimation of causal parameters under interference. However, since there is no formal theory or practical guidance as to how to verify the

[1] In the literature, Forastiere *et al.* (2021) consider an exposure mapping whose range may be heterogeneous across i and n. Although our results would hold with minor modifications even in the presence of heterogeneity in exposure mappings, we let this be beyond the scope of this paper since such a generalization substantially complicates the asymptotic theory. Nevertheless, the common range assumption should not be too restrictive in practice, given that in our framework researchers can arbitrarily specify the form of IEM.

[2] Note that the definition in (2.1) does not imply the uniqueness of correct IEM. For example, if the neighborhood maximum of Z, $T_i = \max\{Z_j : A_{ij} = 1\}$, is a correct IEM for i, so is i's neighborhood average. If there do not exist any spillover effects in the first place, then any IEM is correct.

specification of IEM, in reality, IEMs are generally misspecified.[3] In this case, $\widetilde{y}_i(z, t)$ and $\widetilde{d}_i(z, t)$ are no longer well-defined.

Throughout the paper, following the recent literature on causal inference with interference, we focus on a design-based uncertainty framework where the randomness comes only from $\boldsymbol{Z}$. That is, $\boldsymbol{Z}$ is the only random component in the model, and we treat the potential outcomes, the potential treatments, and the adjacency matrix as non-stochastic components. This design-based approach is suitable for randomized experiments where researchers can design the random assignment mechanism for treatment eligibility or initial recommendations, as in the experiments in Dupas (2014) and Paluck *et al.* (2016). Even in observational studies, the design-based approach should be relevant when we can observe the entire population or most of the finite population (cf. Abadie *et al.*, 2020). It is important to note that we can also view our framework as a full random design conditional on the potential outcomes, the potential treatments, and the adjacency matrix.

Here, as in the standard IV model, we consider the assumption that the IV can affect the outcome only through the treatment, i.e., the exclusion restriction:

Assumption 2.1 (Exclusion restriction)**.** $Y_i(\boldsymbol{d}, \boldsymbol{z}) = Y_i(\boldsymbol{d}, \boldsymbol{z}')$ for all $i \in N_n$, $\boldsymbol{d} \in \mathcal{D}_n$, and $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}_n$.

Under Assumption 2.1, we can reduce the potential outcome when $\boldsymbol{D} = \boldsymbol{d}$ to $Y_i(\boldsymbol{d}) = Y_i(\boldsymbol{d}, \boldsymbol{z})$, and we have $y_i(\boldsymbol{z}) = Y_i(\boldsymbol{D}(\boldsymbol{z}))$. Note that this assumption is not an essentially necessary condition in terms of ITT analysis, but it can greatly improve the causal interpretation of the parameters that we are going to estimate.

Finally in this section, for illustrative purposes, we provide three specific examples that can be effectively analyzed within our model.

Example 2.1. Suppose that the observed outcome is generated by the following linear model:

$$Y_i = \beta_{0i} + \beta_1 D_i + \beta_2 \cdot \mathbf{1}\left\{\sum_{j \neq i} A_{ij} D_j > c\right\},$$

where β_{0i} is an idiosyncratic intercept term, β_1 and β_2 indicate the direct effect and the spillover effect, respectively, and c is a given threshold. Assume that the treatment status of each unit is determined only by her own IV, i.e., $D_i(z_i) = D_i(z_i, \boldsymbol{z}_{-i})$. Then, the potential outcome when $\boldsymbol{Z} = \boldsymbol{z}$ can be written as

$$y_i(\boldsymbol{z}) = \beta_{0i} + \beta_1 D_i(z_i) + \beta_2 \cdot \mathbf{1}\left\{\sum_{j \neq i} A_{ij} D_j(z_j) > c\right\}.$$

A correctly specified IEM is, for example, $T(i, \boldsymbol{Z}, \boldsymbol{A}) = \mathbf{1}\{\sum_{j \neq i} A_{ij} D_j(Z_j) > c\}$ with $\mathcal{T} = \{0, 1\}$. In the literature, this type of exposure mapping is used, for example, in Hong and Raudenbush (2006) and Leung (2021). With this IEM, one can easily find that the direct effect can be obtained by a Wald-type estimand:

$$\beta_1 = \frac{\sum_{i \in S_n} \mathbb{E}[Y_i \mid Z_i = 1, T_i = t] - \sum_{i \in S_n} \mathbb{E}[Y_i \mid Z_i = 0, T_i = t]}{\sum_{i \in S_n} \mathbb{E}[D_i \mid Z_i = 1, T_i = t] - \sum_{i \in S_n} \mathbb{E}[D_i \mid Z_i = 0, T_i = t]},$$

[3] The IEM is said to be *misspecified* if there exist some $i \in N_n$, $z_i \in \{0, 1\}$, $\boldsymbol{z}_{-i}, \boldsymbol{z}'_{-i} \in \{0, 1\}^{n-1}$, and $\boldsymbol{A} \in \mathcal{A}_n$ such that

$$T(i, z_i, \boldsymbol{z}_{-i}, \boldsymbol{A}) = T(i, z_i, \boldsymbol{z}'_{-i}, \boldsymbol{A}) \text{ but } D_i(z_i, \boldsymbol{z}_{-i}) \neq D_i(z_i, \boldsymbol{z}'_{-i}) \text{ or/and } y_i(z_i, \boldsymbol{z}_{-i}) \neq y_i(z_i, \boldsymbol{z}'_{-i}).$$

where S_n is an appropriately chosen subset of N_n. We will show in Theorem 4.1 that the terms on the right-hand side, and hence β_1, can be consistently estimated under certain regularity conditions.

Example 2.2. Assume that Assumption 2.1 holds and that no interference exists in the outcome. For the treatment choice equation, consider the following latent index model:

$$D_i = \mathbf{1}\left\{\gamma_{0i} + \gamma_{1i} Z_i + \gamma_{2i} \cdot \mathbf{1}\left\{\sum_{j \neq i} A_{ij} Z_j > c\right\} > 0\right\},$$

where γ_{0i} is the preference heterogeneity for the treatment, and γ_{1i} and γ_{2i} respectively capture the direct and spillover effect of the IV on the treatment choice. In this situation, the potential outcome when $\boldsymbol{Z} = \boldsymbol{z}$ is given by $y_i(\boldsymbol{z}) = \beta_{0i} + \beta_{1i} D_i(\boldsymbol{z})$ without loss of generality. As such, the model is a simple binary treatment model with potentially many IVs. It is straightforward to find that we can estimate a local average treatment effect (LATE)-type parameter using the two-stage least squares method under a monotonicity condition between D_i and Z_i (e.g., $\gamma_{1i} \geqslant 0$ for all i), without considering the spillover effect in the treatment choice model. If we set $T(i, \boldsymbol{Z}, \boldsymbol{A}) = \mathbf{1}\{\sum_{j \neq i} A_{ij} Z_j > c\}$, this is clearly a correct IEM.

Example 2.3. The following linear-in-means model has been often used in applied studies of peer effects (cf. Bramoullé *et al.*, 2009):

$$Y_i = \beta_{0i} + \beta_1 \frac{\sum_{j \neq i} A_{ij} Y_j}{\sum_{j \neq i} A_{ij}} + \beta_2 D_i,$$

assuming that each unit has at least one link. Letting $\boldsymbol{Y} = (Y_i)_{i \in N_n}$, $\boldsymbol{\beta}_0 = (\beta_{0i})_{i \in N_n}$, and $\boldsymbol{G} = (G_{ij})_{i,j \in N_n}$ with $G_{ij} = A_{ij} / \sum_{j \neq i} A_{ij}$, we can re-write the model in vector-form as $\boldsymbol{Y} = \boldsymbol{\beta}_0 + \beta_1 \boldsymbol{G}\boldsymbol{Y} + \beta_2 \boldsymbol{D}$. If $|\beta_1| < 1$ holds, $I_n - \beta_1 \boldsymbol{G}$ is nonsingular, where I_n is the identity matrix of dimension n, and we have

$$\boldsymbol{y}(\boldsymbol{Z}) = (I_n - \beta_1 \boldsymbol{G})^{-1}(\boldsymbol{\beta}_0 + \beta_2 \boldsymbol{D}(\boldsymbol{Z})).$$

This expression clearly shows that in general the potential outcome $y_i(\boldsymbol{z})$ relates to all z_i's in a nontrivial manner. When $D_i(z_i) = D_i(z_i, \boldsymbol{z}_{-i})$ is true, $T(i, \boldsymbol{Z}, A) = \sum_{j \neq i} G_{ij} y_j(\boldsymbol{Z})$ is a correct IEM specification.

3 Identification

This section discusses the identification of treatment parameters with potentially misspecified IEM. We first study the ITT effects in Section 3.1. We then develop our main identification result for the LADE in Section 3.2. Throughout this section, we assume the following:

Assumption 3.1 (Conditional independence)**.** For all $i \in N_n$, Z_i is conditionally independent of $\boldsymbol{Z}_{-i}$ given T_i.

This assumption restricts the specification of IEM and the joint distribution of $\boldsymbol{Z} = (Z_i, \boldsymbol{Z}_{-i})$. For example, suppose that Z_i's are independent and identically distributed (IID) across i. Then, Assumption 3.1 would be satisfied when Z_i is not a determinant of T_i. Such examples include the instrumental exposures considered in Examples 2.1 and 2.2. Because researchers can select the form of IEM to ensure Assumption 3.1, this assumption may not be too restrictive in practice. Nevertheless, it should be noted that the interpretations of the causal estimands presented below may differ significantly for different IEM specifications.

3.1 Intention-to-treat analysis

In this subsection, we provide the identification results for the ADE and ASE of the IV on the outcome and on the treatment choice. Although the results in this subsection are simple corollaries of those in prior studies (e.g., Aronow and Samii, 2017; Leung, 2021; Sävje, 2021), we describe them in some detail because they form the basis for the identification analysis of the LADE parameter.

To proceed, consider a non-random sub-population $S_n \subseteq N_n$. Throughout the paper, we consider estimating causal parameters specific to this sub-population. For an example of S_n, let $S_n(\delta)$ be the set of units whose degrees are δ: $S_n(\delta) = \{i \in N_n : \sum_{i \neq j} A_{ij} = \delta\}$. In this case, we can examine whether the causal impacts are heterogeneous across individuals with different centrality by comparing the parameter estimates obtained from $S_n(\delta)$ with different δ's.

Define

$$\mu_i^Y(z,t) := \mathbb{E}[Y_i \mid Z_i = z, T_i = t], \qquad \mu_i^D(z,t) := \mathbb{E}[D_i \mid Z_i = z, T_i = t],$$

for $z \in \{0,1\}$ and $t \in \mathcal{T}$. Here, the expectation is taken with respect to the distribution of $\boldsymbol{Z}$. Since these quantities may vary with individuals due to heterogeneity in the potential outcome and potential treatment choice, generally we cannot obtain consistent estimators for them in the design-based approach. Denote their averages over S_n as

$$\bar{\mu}_{S_n}^Y(z,t) := \frac{1}{|S_n|}\sum_{i \in S_n} \mu_i^Y(z,t), \qquad \bar{\mu}_{S_n}^D(z,t) := \frac{1}{|S_n|}\sum_{i \in S_n} \mu_i^D(z,t),$$

where $|S_n|$ is the cardinality of S_n. We will show in Section 4 that $\bar{\mu}_{S_n}^Y(z,t)$ and $\bar{\mu}_{S_n}^D(z,t)$ can be consistently estimated under ANI in our context (see Assumption 4.5) and certain conditions on the network structure.

The ADE of the IV on the outcome and that on the treatment receipt are respectively defined by

$$\mathrm{ADEY}_{S_n}(t) := \bar{\mu}_{S_n}^Y(1,t) - \bar{\mu}_{S_n}^Y(0,t), \qquad \mathrm{ADED}_{S_n}(t) := \bar{\mu}_{S_n}^D(1,t) - \bar{\mu}_{S_n}^D(0,t),$$

for $t \in \mathcal{T}$. Similarly, we define the ASEs by

$$\mathrm{ASEY}_{S_n}(z,t,t') := \bar{\mu}_{S_n}^Y(z,t) - \bar{\mu}_{S_n}^Y(z,t'), \qquad \mathrm{ASED}_{S_n}(z,t,t') := \bar{\mu}_{S_n}^D(z,t) - \bar{\mu}_{S_n}^D(z,t'),$$

for $z \in \{0,1\}$ and $t, t' \in \mathcal{T}$. It should be noted that these quantities are well-defined irrespective of whether the IEM is correctly specified or not. The following proposition presents the causal interpretation of these estimands.

Proposition 3.1. Let $\pi_i(\boldsymbol{z}_{-i}, t) := \Pr[\boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i} \mid T_i = t]$. Under Assumption 3.1, we have

$$\begin{aligned}
\text{ADEY}_{S_n}(t) &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \{y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})\} \pi_i(\boldsymbol{z}_{-i}, t), \\
\text{ADED}_{S_n}(t) &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \{D_i(1, \boldsymbol{z}_{-i}) - D_i(0, \boldsymbol{z}_{-i})\} \pi_i(\boldsymbol{z}_{-i}, t), \\
\text{ASEY}_{S_n}(z, t, t') &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} y_i(z, \boldsymbol{z}_{-i}) \{\pi_i(\boldsymbol{z}_{-i}, t) - \pi_i(\boldsymbol{z}_{-i}, t')\}, \\
\text{ASED}_{S_n}(z, t, t') &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} D_i(z, \boldsymbol{z}_{-i}) \{\pi_i(\boldsymbol{z}_{-i}, t) - \pi_i(\boldsymbol{z}_{-i}, t')\}.
\end{aligned}$$

Proposition 3.1 shows that the $\text{ADEY}_{S_n}(t)$ and $\text{ADED}_{S_n}(t)$ have a causal interpretation as the weighted average of $y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})$ and as that of $D_i(1, \boldsymbol{z}_{-i}) - D_i(0, \boldsymbol{z}_{-i})$, respectively, with the weight equal to $\pi_i(\boldsymbol{z}_{-i}, t)$. If we additionally impose Assumption 2.1, the result for $\text{ADEY}_{S_n}(t)$ can be further well interpreted. To see this, let $D_j(Z_i = z_i, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})$ be the potential treatment status of unit $j \neq i$ when $Z_i = z_i$ and $\boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}$, and let $\boldsymbol{D}_{-i}(Z_i = z_i, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}) = (D_j(Z_i = z_i, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}))_{j \neq i}$. By the definition of $y_i(z_i, \boldsymbol{z}_{-i})$ and Assumption 2.1, we can observe that

$$\begin{aligned}
y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i}) = &\underbrace{Y_i(D_i(1, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) - Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}))}_{\text{direct effect of IV}} \\
&+ \underbrace{Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) - Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}))}_{\text{spillover effect of IV}}.
\end{aligned}$$

That is, $y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})$ comprises of the direct effect of changing i's own treatment status from $D_i(0, \boldsymbol{z}_{-i})$ to $D_i(1, \boldsymbol{z}_{-i})$ and the spillover effect by changing the others' treatments from $\boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})$ to $\boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})$. Hence, Proposition 3.1 can be read as that $\text{ADEY}_{S_n}(t)$ consists of the sum of the ADE from own IV and the ASE caused by changing the unit's own IV.

Given Proposition 3.1, the results for the ASE would be less interpretable than those for the ADE. This difficulty in interpreting ASEs is mainly due to the possibility of misspecification of the IEM. That said, even when the IEM is not properly specified, ASE contains beneficial information regarding spillover effects. A non-zero estimate of ASE indicates the presence of some form of interference.

Proposition 3.1 is also useful for interpreting estimates of ADE or ASE obtained from different forms of IEM. For example, suppose that two IEMs T and T' generate the same estimates of ADEY at t and t', respectively:

$$\begin{aligned}
0 &= \text{ADEY}_{S_n}(t)|_{\text{IEM}=T} - \text{ADEY}_{S_n}(t')|_{\text{IEM}=T'} \\
&= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \{y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})\} \cdot \{\pi_i(\boldsymbol{z}_{-i}, t) - \pi'_i(\boldsymbol{z}_{-i}, t')\},
\end{aligned}$$

This is possible for example (i) when the treatment effect $y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})$ is homogeneous with respect to $\boldsymbol{z}_{-i}$ for all individuals (i.e., no spillovers), or (ii) when $\{T_i = t\}$ and $\{T'_i = t'\}$ are essentially the same conditions such that $\pi_i(\boldsymbol{z}_{-i}, t) - \pi_i'(\boldsymbol{z}_{-i}, t') = 0$. Note that (i) is testable since if (i) is true the above equality must hold for any IEMs.

Remark 3.1 (Correctly specified IEM)**.** For now, suppose that the IEM T satisfies (2.1). In this situation, the potential outcome and the potential treatment status given $Z_i = z$ and $T_i = t$ are $\widetilde{y}_i(z, t)$ and $\widetilde{d}_i(z, t)$, respectively, and we have $\mu_i^Y(z, t) = \widetilde{y}_i(z, t)$ and $\mu_i^D(z, t) = \widetilde{d}_i(z, t)$. Then, it is straightforward to see that $\text{ADEY}_{S_n}(t) = |S_n|^{-1} \sum_{i \in S_n} [\widetilde{y}_i(1, t) - \widetilde{y}_i(0, t)]$, and a similar result applies to the other parameters as well. Assumption 3.1 is unnecessary for this result to hold.

3.2 Local average direct effect

In this subsection, we present our main identification result for the LADE. Throughout this subsection, we maintain Assumption 2.1. Firstly, we extend the notion of *compliers* (Angrist *et al.*, 1996; Imai *et al.*, 2021) to our setting. Let

$$\mathcal{C}_i(\boldsymbol{z}_{-i}) := \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) = 1, D_i(0, \boldsymbol{z}_{-i}) = 0\}$$

be an indicator for being a complier who takes the treatment only when $Z_i = 1$ conditional on $\boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}$. Thus, the compliance status may depend on the assignment of IVs to the others. Denote the realized compliance status of unit i as $\mathcal{C}_i := \mathcal{C}_i(\boldsymbol{Z}_{-i})$. The expected compliance status conditional on $T_i = t$ is given by

$$\mathbb{E}[\mathcal{C}_i \mid T_i = t] = \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathcal{C}_i(\boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t).$$

The LADE is defined by the weighted average of $y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})$ over the compliers:

$$\text{LADE}_{S_n}(t) := \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \{y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i})\} \frac{\mathcal{C}_i(\boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t)}{\sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathcal{C}_i(\boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t)},$$

provided that the denominator $\sum_{i \in S_n} \mathbb{E}[\mathcal{C}_i \mid T_i = t]$ is non-zero. By the same decomposition as above, for a complier i such that $\mathcal{C}_i(\boldsymbol{z}_{-i}) = 1$,

$$\begin{aligned} y_i(1, \boldsymbol{z}_{-i}) - y_i(0, \boldsymbol{z}_{-i}) &= \underbrace{Y_i(1, \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) - Y_i(0, \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}))}_{\text{direct effect of treatment}} \\ &\quad + \underbrace{Y_i(0, \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) - Y_i(0, \boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}))}_{\text{spillover effect of IV}}. \end{aligned}$$

As such, the LADE parameter captures the sum of the average direct treatment effect and the ASE caused by changing the unit's own IV, for the compliers.

In what follows, we introduce a set of sufficient conditions for identifying the LADE. The following two conditions are analogous to the IV relevance condition and the monotonicity condition for the standard LATE estimation without interference.

Assumption 3.2 (Relevance)**.** There exists a constant $c > 0$ (which may depend on T and t) such that $|S_n|^{-1} \sum_{i \in S_n} \mathbb{E}[\mathcal{C}_i \mid T_i = t] \geqslant c$.

Assumption 3.3 (Monotonicity)**.** $D_i(1, \boldsymbol{z}_{-i}) \geqslant D_i(0, \boldsymbol{z}_{-i})$ for all $i \in S_n$ and $\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}$ such that $\pi_i(\boldsymbol{z}_{-i}, t) > 0$.

Assumption 3.2 states that there is a non-negligible proportion of units among those with $T_i = t$ whose treatment status is positively affected by their assigned IV. The assumption is necessary to well-define the LADE parameter. Assumption 3.3 requires that there do not exist *defiers*, whose treatment status is negatively affected by the instrument. This assumption limits the heterogeneity in treatment choice in that the response to the IV must be non-negative for all units. For instance, the treatment choice equation in Example 2.2 satisfies the monotonicity condition if $\gamma_{1i} \geqslant 0$ for all i. Under Assumptions 3.2 and 3.3, each individual can be classified into one of the three latent types: compliers; *always takers* (those who always take the treatment); *never takers* (those who never take the treatment).

Unlike conventional identification results without interference, the set of the exclusion restriction, the relevance condition, and the monotonicity condition does not suffice to identify the LADE parameter. This is because we need to account for two potential interference channels at the same time: one is the spillover effect of the IV on the treatment receipt, and the other is the spillover effect of the treatment on the outcome. As in the conventional method, we use the variation in the value of IV to identify the LADE, but in the present situation, the effect of shifting IV can be amplified in two steps by the two different spillovers. Therefore, to facilitate the identification of the LADE parameter, some additional restriction on the interference structure is needed. In this study, similar to Imai *et al.* (2021), we require the potential outcome $y_i(z_i, \boldsymbol{z}_{-i})$ of noncompliers to be insensitive to their own instrumental value z_i.

Assumption 3.4 (Restricted interference)**.** For all $i \in S_n$ and $\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}$ such that $\pi_i(\boldsymbol{z}_{-i}, t) > 0$, $y_i(1, \boldsymbol{z}_{-i}) = y_i(0, \boldsymbol{z}_{-i})$ holds whenever $D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i})$.

Here, we provide three empirically relevant sufficient conditions for this assumption. The first condition is no spillovers between IV and the treatment choice:

$$D_i(z_i, \boldsymbol{z}_{-i}) = D_i(z_i, \boldsymbol{z}'_{-i}) \quad \text{for any } z_i \in \{0,1\} \text{ and } \boldsymbol{z}_{-i}, \boldsymbol{z}'_{-i} \in \{0,1\}^{n-1}. \tag{3.1}$$

This corresponds to the personalized encouragement assumption of Kang and Imbens (2016), which states that an incentive to take treatment must be personalized to everyone. Under this condition, we can define the potential treatment status as $D_i(z_i) = D_i(z_i, \boldsymbol{z}_{-i})$. Then, the potential outcome satisfies $y_i(z_i, \boldsymbol{z}_{-i}) = Y_i(D_i(z_i), (D_j(z_j))_{j \neq i})$, implying Assumption 3.4.

The second situation in which Assumption 3.4 holds is when there is no treatment spillover effect on the outcome; that is,

$$Y_i(d_i, \boldsymbol{d}_{-i}) = Y_i(d_i, \boldsymbol{d}'_{-i}) \quad \text{for any } d_i \in \{0,1\} \text{ and } \boldsymbol{d}_{-i}, \boldsymbol{d}'_{-i} \in \{0,1\}^{n-1}. \tag{3.2}$$

Then, we may write the potential outcome given $D_i = d_i$ as $Y_i(d_i)$. It is easy to see that (3.2) implies Assumption 3.4.

The third sufficient condition for Assumption 3.4 is that the IV of any noncomplier does not affect the treatment status of all other units; specifically, for any $\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}$,

$$\boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}) = \boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}) \quad \text{whenever } D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i}). \tag{3.3}$$

If this condition holds, the potential outcome of unit i with $D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i})$ satisfies

$$\begin{aligned} y_i(1, \boldsymbol{z}_{-i}) &= Y_i(D_i(1, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) \\ &= Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) = y_i(0, \boldsymbol{z}_{-i}), \end{aligned}$$

which implies Assumption 3.4.

As such, the interpretation of the LADE parameter can be different depending on which sufficient condition the researcher considers for Assumption 3.4 to hold. For example, under (3.1) or (3.2),

$$Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) - Y_i(D_i(0, \boldsymbol{z}_{-i}), \boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) = 0.$$

Therefore, the LADE is identical to solely the ADE of the treatment on the outcome for the compliers. On the contrary, (3.3) does not restrict the interference structure for the compliers so that the IV of a complier may affect the treatment status of others. In this case, as discussed above, the LADE parameter is considered as the sum of the direct effect and the spillover effect for the compliers.

The following theorem shows that the LADE parameter can be identified by a Wald-type estimand.

Theorem 3.1. Under Assumptions 3.1 – 3.4, it holds that

$$\mathrm{LADE}_{S_n}(t) = \frac{\mathrm{ADEY}_{S_n}(t)}{\mathrm{ADED}_{S_n}(t)}. \tag{3.4}$$

Remark 3.2 (Average noncompliance rate)**.** Let

$$\mathcal{A}_i(\boldsymbol{z}_{-i}) := \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i}) = 1\} \text{ and } \mathcal{N}_i(\boldsymbol{z}_{-i}) := \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i}) = 0\}$$

denote the indicators for being an always taker and a never taker, respectively, conditional on $\boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}$. Their realized states are $\mathcal{A}_i := \mathcal{A}_i(\boldsymbol{Z}_{-i})$ and $\mathcal{N}_i := \mathcal{N}_i(\boldsymbol{Z}_{-i})$. Under Assumptions 3.1 and 3.3, similar arguments to the proof of Theorem 3.1 can show that

$$\begin{aligned} \frac{1}{|S_n|} \sum_{i \in S_n} \mathbb{E}[D_i \mid Z_i = 0, T_i = t] &= \frac{1}{|S_n|} \sum_{i \in S_n} \mathbb{E}[\mathcal{A}_i \mid T_i = t], \\ \frac{1}{|S_n|} \sum_{i \in S_n} \mathbb{E}[1 - D_i \mid Z_i = 1, T_i = t] &= \frac{1}{|S_n|} \sum_{i \in S_n} \mathbb{E}[\mathcal{N}_i \mid T_i = t]. \end{aligned}$$

Thus, we can measure the average noncompliance status by computing the left-hand sides of these equalities.

Remark 3.3 (Testable implication of (3.1))**.** Under condition (3.1), Proposition 3.1 implies that $\mathrm{ASED}_{S_n}(z, t, t') = 0$ for all $z \in \{0, 1\}$ and $t, t' \in \mathcal{T}$. Thus, if $\mathrm{ASED}_{S_n}(z, t, t') \neq 0$ is observed, it indicates there is a violation of condition (3.1); however, the converse is generally not true. Of course, if the specification of IEM is correct, we can support (3.1) by confirming $\text{ASED}_{S_n}(z, t, t') = 0$ for all z, t, and t'.

Remark 3.4 (Testable implication of (3.2))**.** Denote $m_i^{(1)}(y, z, t) := \mathbb{E}[\mathbf{1}\{Y_i \leqslant y\} D_i \mid Z_i = z, T_i = t]$ for $y \in \mathbb{R}$, $z \in \{0, 1\}$, and $t \in \mathcal{T}$. Condition (3.2) implies that $m_i^{(1)}(y, z, t) = \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{Y_i(1) \leqslant y\} D_i(z, \boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t)$ by Assumption 3.1. As a result, we have $m_i^{(1)}(y, 1, t) - m_i^{(1)}(y, 0, t) \geqslant 0$ under Assumption 3.3. Note that this inequality might not hold without condition (3.2). Similarly, letting $m_i^{(0)}(y, z, t) := \mathbb{E}[\mathbf{1}\{Y_i \leqslant y\}(1 - D_i) \mid Z_i = z, T_i = t]$, we can show that $m_i^{(0)}(y, 1, t) - m_i^{(0)}(y, 0, t) \leqslant 0$. Note that although these inequalities cannot be directly tested for each i, we can check whether they hold or not on average for some sub-samples.

Remark 3.5 (Testable implication of (3.3))**.** Let $g_i : \{0, 1\}^{n-1} \to \mathbb{R}_+$ be a known non-negative function of $\boldsymbol{D}_{-i}$. Then, we can show that

$$\mathbb{E}[D_i g_i(\boldsymbol{D}_{-i}) \mid Z_i = 0, T_i = t] = \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathcal{A}_i(\boldsymbol{z}_{-i}) g_i(\boldsymbol{D}_{-i}(Z_i = 0, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) \pi_i(\boldsymbol{z}_{-i}, t)$$

by Assumptions 3.1 and 3.3. Similarly,

$$\mathbb{E}[D_i g_i(\boldsymbol{D}_{-i}) \mid Z_i = 1, T_i = t] = \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \{\mathcal{C}_i(\boldsymbol{z}_{-i}) + \mathcal{A}_i(\boldsymbol{z}_{-i})\} g_i(\boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) \pi_i(\boldsymbol{z}_{-i}, t).$$

Thus, condition (3.3) leads to

$$\begin{aligned} &\mathbb{E}[D_i g_i(\boldsymbol{D}_{-i}) \mid Z_i = 1, T_i = t] - \mathbb{E}[D_i g_i(\boldsymbol{D}_{-i}) \mid Z_i = 0, T_i = t] \\ &\quad = \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathcal{C}_i(\boldsymbol{z}_{-i}) g_i(\boldsymbol{D}_{-i}(Z_i = 1, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i})) \pi_i(\boldsymbol{z}_{-i}, t) \geqslant 0. \end{aligned}$$

A similar argument shows that

$$\mathbb{E}[(1 - D_i) g_i(\boldsymbol{D}_{-i}) \mid Z_i = 0, T_i = t] - \mathbb{E}[(1 - D_i) g_i(\boldsymbol{D}_{-i}) \mid Z_i = 1, T_i = t] \geqslant 0.$$

Then, these inequalities provide testable implications of (3.3).

4 Estimation and Asymptotic Theory

In this section, we discuss the nonparametric estimation of the causal parameters presented in the previous section. The estimation procedures are discussed in Section 4.1, and Section 4.2 presents their asymptotic properties. In Section 4.3, we provide statistical inference methods.

4.1 Estimators

We consider the following data generating process (DGP):

Assumption 4.1 (DGP)**.**

(i) $\{Z_i\}_{i \in N_n}$ are mutually independent.

(ii) $\{(T_i, Z_i)\}_{i \in S_n}$ are identically distributed across $i \in S_n$.

(iii) $\mathcal{T}$ is a finite subset of $\mathbb{R}^{\dim(T)}$.

Assumption 4.1(i) would be reasonable for many empirical situations. For example, the assumption is relevant to a randomized experiment where the treatment eligibility is assigned to each unit with some probability $q \in (0,1)$ (i.e., a Bernoulli trial). Another example where the assumption can hold is an observational study in which the IV of each unit is determined separately from the other units. Assumption 4.1(ii) can be justified by appropriately choosing IEM T and sub-population S_n. For example, this assumption holds when $T_i = \mathbf{1}\{\sum_{j \neq i} A_{ij} Z_j > c\}$ and $S_n = \{i \in N_n : \sum_{j \neq i} A_{ij} = \delta\}$ for some c and δ, provided that $\{Z_i\}_{i \in N_n}$ are IID. We require this assumption in order to construct a consistent estimator for the generalized propensity score $\Pr[Z_i = z, T_i = t \mid i \in S_n]$. Lastly, Assumption 4.1(iii) restricts that the IEM takes finite values, which facilitates our asymptotic analysis.

Under Assumption 4.1, we can write $p_{S_n}(z,t) := \Pr[Z_i = z, T_i = t \mid i \in S_n]$. In the following, the condition $i \in S_n$ in the expectation is suppressed for notational simplicity. We estimate $p_{S_n}(z,t)$ by

$$\widehat{p}_{S_n}(z,t) := \frac{1}{|S_n|} \sum_{i \in S_n} \mathbf{1}\{Z_i = z, T_i = t\}.$$

Then, $\bar{\mu}^Y_{S_n}(z,t)$ and $\bar{\mu}^D_{S_n}(z,t)$ can be estimated respectively by

$$\widehat{\mu}^Y_{S_n}(z,t) := \frac{1}{|S_n|} \sum_{i \in S_n} \frac{Y_i \mathbf{1}\{Z_i = z, T_i = t\}}{\widehat{p}_{S_n}(z,t)}, \qquad \widehat{\mu}^D_{S_n}(z,t) := \frac{1}{|S_n|} \sum_{i \in S_n} \frac{D_i \mathbf{1}\{Z_i = z, T_i = t\}}{\widehat{p}_{S_n}(z,t)}.$$

Given these estimators, we compute

$$\widehat{\text{ADEY}}_{S_n}(t) := \widehat{\mu}^Y_{S_n}(1,t) - \widehat{\mu}^Y_{S_n}(0,t), \qquad \widehat{\text{ASEY}}_{S_n}(z,t,t') := \widehat{\mu}^Y_{S_n}(z,t) - \widehat{\mu}^Y_{S_n}(z,t'),$$

$$\widehat{\text{ADED}}_{S_n}(t) := \widehat{\mu}^D_{S_n}(1,t) - \widehat{\mu}^D_{S_n}(0,t), \qquad \widehat{\text{ASED}}_{S_n}(z,t,t') := \widehat{\mu}^D_{S_n}(z,t) - \widehat{\mu}^D_{S_n}(z,t').$$

The identification result in Theorem 3.1 leads to the following estimator for the LADE:

$$\widehat{\text{LADE}}_{S_n}(t) := \frac{\widehat{\text{ADEY}}_{S_n}(t)}{\widehat{\text{ADED}}_{S_n}(t)}.$$

Remark 4.1 (Unbiased estimation of ADEs and ASEs)**.** The proposed estimators may have finite sample bias due to the estimation of $p_{S_n}(z,t)$. Meanwhile, in an experimental situation where $p_{S_n}(z,t)$ can be directly computed from the known distribution of $\boldsymbol{Z}$, we can achieve unbiased estimation of ADEs and ASEs. Then, we can estimate $\bar{\mu}^Y_{S_n}(z,t)$ and $\bar{\mu}^D_{S_n}(z,t)$ respectively by

$$\check{\mu}^Y_{S_n}(z,t) := \frac{1}{|S_n|} \sum_{i \in S_n} \frac{Y_i \mathbf{1}\{Z_i = z, T_i = t\}}{p_{S_n}(z,t)}, \qquad \check{\mu}^D_{S_n}(z,t) := \frac{1}{|S_n|} \sum_{i \in S_n} \frac{D_i \mathbf{1}\{Z_i = z, T_i = t\}}{p_{S_n}(z,t)} \tag{4.1}$$

without biases.

4.2 Asymptotic properties

In this section, we prove the consistency and asymptotic normality of the proposed estimators. In the next two assumptions, we require the potential outcome $y_i(\boldsymbol{z})$ and the generalized propensity score $p_{S_n}(z,t)$ to be uniformly bounded by some constants.

Assumption 4.2 (Bounded outcome)**.** There exists a constant $\bar{y}$ such that $|y_i(\boldsymbol{z})| \leqslant \bar{y} < \infty$ for all $i \in S_n$ and $\boldsymbol{z} \in \mathcal{Z}_n$.

Assumption 4.3 (Overlap)**.** There exist constants $\underline{p}$ and $\bar{p}$ such that $p_{S_n}(z,t) \in [\underline{p}, \bar{p}] \subset (0,1)$ for all $z \in \{0,1\}$ and $t \in \mathcal{T}$.

These requirements are popular in the literature but restrict the DGP. In particular, Assumption 4.3 depends on the specification of IEM T, the choice of sub-population S_n, the distribution of $\boldsymbol{Z}$, and the structure of network $\boldsymbol{A}$. For example, when $T_i = \mathbf{1}\{\sum_{j\neq i} A_{ij} Z_j > 0\}$ and $S_n = \{i \in N_n : \sum_{j\neq i} A_{ij} = \delta\}$, this assumption is violated if every unit in S_n has at least one direct neighborhood with $Z_j = 1$.

The next assumption calls for some additional notations. Denote the path distance (defined on the whole population N_n) between units i and j as $\ell_{\boldsymbol{A}}(i,j)$.[4] For a non-negative integer $s \geqslant 0$, let $N_{\boldsymbol{A}}(i,s) := \{j \in N_n : \ell_{\boldsymbol{A}}(i,j) \leqslant s\}$ be the set of units within s distance from unit i; namely, unit i's s-neighborhood. Note that $i \in N_{\boldsymbol{A}}(i,s)$ for all $s \geqslant 0$. We write the sub-vector of $\boldsymbol{z} \in \mathcal{Z}_n$ restricted on $N_{\boldsymbol{A}}(i,s)$ as $\boldsymbol{z}_{N_{\boldsymbol{A}}(i,s)} := (z_j)_{j\in N_{\boldsymbol{A}}(i,s)}$. Similarly, let $\boldsymbol{A}_{N_{\boldsymbol{A}}(i,s)} = (A_{kl})_{k,l\in N_{\boldsymbol{A}}(i,s)}$ denote the sub-matrix of $\boldsymbol{A}$ restricted on $N_{\boldsymbol{A}}(i,s)$.

Assumption 4.4 (IEM)**.** There exists a known positive integer $K \in \mathbb{N}$ such that, for all $i \in S_n$, $\boldsymbol{A}, \boldsymbol{A}' \in \mathcal{A}_n$, and $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{Z}_n$,

$$N_{\boldsymbol{A}}(i,K) = N_{\boldsymbol{A}'}(i,K),\ \boldsymbol{A}_{N_{\boldsymbol{A}}(i,K)} = \boldsymbol{A}'_{N_{\boldsymbol{A}'}(i,K)},\text{ and } \boldsymbol{z}_{N_{\boldsymbol{A}}(i,K)} = \boldsymbol{z}'_{N_{\boldsymbol{A}'}(i,K)} \quad\Longrightarrow\quad T(i,\boldsymbol{z},\boldsymbol{A}) = T(i,\boldsymbol{z}',\boldsymbol{A}').$$

The assumption states that the instrumental exposure of each unit depends only on the unit's own K-neighborhood. This would be a mild requirement that most IEMs of practical interest should satisfy. For instance, in our empirical example, we consider $T_i = \mathbf{1}\{\sum_{j\neq i} A_{ij} Z_j > 0\}$ and $T_i = \mathbf{1}\{\sum_{j\neq i} A_{ij} D_j > 0\}$ for which $K = 1$.

Next, we introduce the concept of ANI, which is recently proposed in Leung (2021). Let $N^{c}_{\boldsymbol{A}}(i,s) := N_n \backslash N_{\boldsymbol{A}}(i,s)$ denote the set of units who are more than distance s away from i. Writing $\boldsymbol{Z}'$ as an independent copy of $\boldsymbol{Z}$, we define $\boldsymbol{Z}_i^{(s)} := (\boldsymbol{Z}_{N_{\boldsymbol{A}}(i,s)}, \boldsymbol{Z}'_{N^{c}_{\boldsymbol{A}}(i,s)})$ by combining the sub-vector of $\boldsymbol{Z}$ on $N_{\boldsymbol{A}}(i,s)$ and that of $\boldsymbol{Z}'$ on $N^{c}_{\boldsymbol{A}}(i,s)$. Denote

$$\theta_{n,s} := \max\left\{ \max_{i\in S_n} \mathbb{E}\,|y_i(\boldsymbol{Z}) - y_i(\boldsymbol{Z}_i^{(s)})|,\ \ \max_{i\in S_n} \mathbb{E}\,|D_i(\boldsymbol{Z}) - D_i(\boldsymbol{Z}_i^{(s)})| \right\}.$$

This quantity measures the intensity of interference with distant units that are at least s distance away. Note that $\theta_{n,s}$ is bounded uniformly in n and $s \geqslant 0$ by Assumption 4.2.

Assumption 4.5 (ANI)**.** $\sup_{n\in\mathbb{N}} \theta_{n,s} \to 0$ as $s \to \infty$.

[4] A *path* between i and j is a sequence of links $A_{k_1k_2} = A_{k_2k_3} = \cdots = A_{k_{m-1}k_m} = 1$, where $k_1 = i$ and $k_m = j$. The *length* of this path is $m-1$. The *path distance* between i and j is the length of the shortest path between them. As convention, we define the path distance between i and j as ∞ when no path exists and 0 if $i = j$.

The ANI assumption says that spillover effects from units that are sufficiently far away should be sufficiently small. In particular, those not connected with unit i (i.e., j's with $\ell_{\boldsymbol{A}}(i,j) = \infty$) do not affect the outcome and treatment response of i. Thus, assuming ANI would be reasonable in practical situations where only nearby people can affect one's decision-making. Note that ANI is much weaker than the commonly used assumption of group-wise interference such that $\theta_{n,K} = 0$ for some K, in that it allows interaction with units at any distance as in Example 2.3.

Let $S_{\boldsymbol{A}}^{\partial}(i,s) := \{j \in S_n : \ell_{\boldsymbol{A}}(i,j) = s\}$ be the subset of S_n that are exactly at distance s from unit $i \in S_n$. We denote its k-th sample moment as $M_{S_n}^{\partial}(s;k) := |S_n|^{-1} \sum_{i \in S_n} |S_{\boldsymbol{A}}^{\partial}(i,s)|^k$. When $k = 1$, we write $M_{S_n}^{\partial}(s) = M_{S_n}^{\partial}(s;1)$. Further, define

$$\widetilde{\theta}_{n,s} := \begin{cases} \theta_{n,\lfloor s/2 \rfloor} & \text{for } s > 2\max\{K,1\} \\ 1 & \text{otherwise} \end{cases}, \tag{4.2}$$

where $\lfloor \cdot \rfloor$ indicates the floor function.

Assumption 4.6 (Weak dependence 1)**.**

(i) $\max_{1 \leqslant s \leqslant 2K} M_{S_n}^{\partial}(s) = O(1)$, where K is as given in Assumption 4.4.

(ii) $|S_n|^{-1} \sum_{s=1}^{n-1} M_{S_n}^{\partial}(s) \widetilde{\theta}_{n,s} = o(1)$.

Assumption 4.6(i) rules out that there are a non-negligible proportion of units whose $2K$ neighborhoods in S_n grow to infinite as n increases. For example, this assumption is violated if the network is a complete graph. Assumption 4.6(ii) is analogous to Assumption 5 of Leung (2021) and Assumption 3.2 of Kojevnikov *et al.* (2021). This assumption restricts the rate of convergence of $\widetilde{\theta}_{n,s}$ to zero as $s \to \infty$. For example, consider a ring network where every unit connects only to the two adjacent units. In this case, we can see that $M_{S_n}^{\partial}(s) \leqslant 2$ for all s, and Assumption 4.6(ii) is reduced to the condition $|S_n|^{-1} \sum_{s=1}^{n-1} \widetilde{\theta}_{n,s} = o(1)$.

The following theorem establishes the consistency of the proposed estimators. We omit its proof because it is straightforward from Lemma B.2.

Theorem 4.1. Suppose that Assumptions 4.1 – 4.6 hold. Then, if $|S_n| \to \infty$, we have

$$\text{(i)} \quad \widehat{\text{ADEY}}_{S_n}(t) - \text{ADEY}_{S_n}(t) \xrightarrow{p} 0,$$

$$\text{(ii)} \quad \widehat{\text{ADED}}_{S_n}(t) - \text{ADED}_{S_n}(t) \xrightarrow{p} 0.$$

Additionally, if Assumptions 3.1–3.4 hold, we have

$$\text{(iii)} \quad \widehat{\text{LADE}}_{S_n}(t) - \text{LADE}_{S_n}(t) \xrightarrow{p} 0.$$

Remark 4.2 (Rate of convergence)**.** In view of the proof of Lemma B.2, we can find that the convergence rates of the proposed estimators are determined by the convergence rate given in Assumption 4.6(ii). In particular,

$\sqrt{|S_n|}$-consistency can be achieved if Assumption 4.6(ii) is strengthened to

$$\sum_{s=1}^{n-1} M^{\partial}_{S_n}(s)\widetilde{\theta}_{n,s} = O(1). \tag{4.3}$$

Next, we investigate the asymptotic distributions of our estimators. For each $i \in S_n$, let

$$V_i^{\text{ADEY}} := W_i^Y - \frac{\bar{\mu}^Y_{S_n}(1,t)}{p_{S_n}(1,t)} W_i^Z + \frac{\bar{\mu}^Y_{S_n}(0,t)}{p_{S_n}(0,t)} W_i^{1-Z}, \qquad V_i^{\text{ADED}} := W_i^D - \frac{\bar{\mu}^D_{S_n}(1,t)}{p_{S_n}(1,t)} W_i^Z + \frac{\bar{\mu}^D_{S_n}(0,t)}{p_{S_n}(0,t)} W_i^{1-Z},$$
$$V_i^{\text{LADE}} := \frac{1}{\text{ADED}_{S_n}(t)} V_i^{\text{ADEY}} - \frac{\text{ADEY}_{S_n}(t)}{[\text{ADED}_{S_n}(t)]^2} V_i^{\text{ADED}},$$

where

$$\begin{aligned} W_i^Y &:= Y_i \left[\frac{\mathbf{1}\{Z_i = 1, T_i = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i = 0, T_i = t\}}{p_{S_n}(0,t)} \right], \\ W_i^D &:= D_i \left[\frac{\mathbf{1}\{Z_i = 1, T_i = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i = 0, T_i = t\}}{p_{S_n}(0,t)} \right], \\ W_i^Z &:= \mathbf{1}\{Z_i = 1, T_i = t\}, \qquad W_i^{1-Z} := \mathbf{1}\{Z_i = 0, T_i = t\}. \end{aligned} \tag{4.4}$$

In the proof of the theorem presented below, we will show that the asymptotic distribution of $\widehat{\text{ADEY}}_{S_n}(t) - \text{ADEY}_{S_n}(t)$ can be obtained by that of $|S_n|^{-1} \sum_{i \in S_n} (V_i^{\text{ADEY}} - \mathbb{E}[V_i^{\text{ADEY}}])$ (see (A.2)). Similar results hold for the other cases. Let

$$\left(\sigma^{\text{ADEY}}_{S_n}\right)^2 := \text{Var}\left[\frac{1}{\sqrt{|S_n|}} \sum_{i \in S_n} V_i^{\text{ADEY}} \right], \qquad \left(\sigma^{\text{ADED}}_{S_n}\right)^2 := \text{Var}\left[\frac{1}{\sqrt{|S_n|}} \sum_{i \in S_n} V_i^{\text{ADED}} \right],$$
$$\left(\sigma^{\text{LADE}}_{S_n}\right)^2 := \text{Var}\left[\frac{1}{\sqrt{|S_n|}} \sum_{i \in S_n} V_i^{\text{LADE}} \right].$$

Note that to achieve $\sqrt{|S_n|}$-consistent estimation, these variances are to be bounded.

To derive the asymptotic distributions, we employ the central limit theorem (CLT) for ψ-weakly dependent processes in Kojevnikov *et al.* (2021) (see Definition B.1 for the notion of ψ-weak dependence). Under Assumptions 4.1–4.5, for each $V_i = V_i^{\text{ADEY}}$, V_i^{ADED}, or V_i^{LADE}, we show that $\{V_i\}_{i \in S_n}$ is a ψ-weakly dependent process with the dependence coefficients $\{\widetilde{\theta}_{n,s}\}_{s \geqslant 0}$. Then, we can apply their CLT to our context if we introduce additional restrictions on the network structure. Let $S_{\boldsymbol{A}}(i,s) := \{j \in S_n : \ell_{\boldsymbol{A}}(i,j) \leqslant s\}$ and

$$\Delta_{S_n}(s,m;k) := \frac{1}{|S_n|} \sum_{i \in S_n} \max_{j \in S^{\partial}_{\boldsymbol{A}}(i,s)} |S_{\boldsymbol{A}}(i,m) \backslash S_{\boldsymbol{A}}(j,s-1)|^k,$$

where we take $S_{\boldsymbol{A}}(j,s-1) = \varnothing$ if $s = 0$. This is the k-th sample moment of the maximum number (over j's at distance s from i) of units who are within distance m from i but at least distance s apart from j. In general,

$\Delta_{S_n}(s, m; k)$ increases as m becomes larger, but decreases to zero as s grows. In addition, we define

$$c_{S_n}(s, m; k) := \inf_{\alpha>1} [\Delta_{S_n}(s, m; k\alpha)]^{\frac{1}{\alpha}} \left[M^{\partial}_{S_n}\left(s; \frac{\alpha}{\alpha-1}\right)\right]^{1-\frac{1}{\alpha}}.$$

This is a measure of the denseness of the network, which plays an important role in establishing the CLT for ψ-weakly dependent processes.

Assumption 4.7 (Weak dependence 2)**.** For each $\sigma_{S_n} = \sigma^{\text{ADEY}}_{S_n}$, $\sigma^{\text{ADED}}_{S_n}$, and $\sigma^{\text{LADE}}_{S_n}$, there exist some positive sequence $m_n \to \infty$ and a constant $0 < \varepsilon < 1$ such that for each $k \in \{1, 2\}$

$$\text{(i)} \quad \frac{1}{|S_n|^{k/2}\sigma^{2+k}_{S_n}} \sum_{s=0}^{n-1} c_{S_n}(s, m_n; k)\widetilde{\theta}^{1-\varepsilon}_{n,s} \to 0,$$

$$\text{(ii)} \quad \frac{|S_n|^{k/2}}{\sigma^k_{S_n}} \widetilde{\theta}^{1-\varepsilon}_{n,m_n} \to 0.$$

This assumption corresponds to Assumption 3.4 of Kojevnikov *et al.* (2021).[5] Note that the assumption restricts not only the network structure but also our choice of sub-population S_n. In particular, when σ_{S_n} is bounded by zero and infinity uniformly in n, Assumption 4.7 can be reduced to (i) $|S_n|^{-k/2} \sum_{s=0}^{n-1} c_{S_n}(s, m_n; k)\widetilde{\theta}^{1-\varepsilon}_{n,s} \to 0$ and (ii) $|S_n|^{k/2}\widetilde{\theta}^{1-\varepsilon}_{n,m_n} \to 0$.

The following theorem shows that the proposed estimators are asymptotically normal.

Theorem 4.2. Suppose that Assumptions 4.1 – 4.7 hold. Then, if $|S_n| \to \infty$, we have

$$\text{(i)} \quad \frac{\sqrt{|S_n|}\left(\widehat{\text{ADEY}}_{S_n}(t) - \text{ADEY}_{S_n}(t)\right)}{\sigma^{\text{ADEY}}_{S_n}} \xrightarrow{d} \text{Normal}(0, 1),$$

$$\text{(ii)} \quad \frac{\sqrt{|S_n|}\left(\widehat{\text{ADED}}_{S_n}(t) - \text{ADED}_{S_n}(t)\right)}{\sigma^{\text{ADED}}_{S_n}} \xrightarrow{d} \text{Normal}(0, 1),$$

provided that $(\sigma^{\text{ADEY}}_{S_n})^{-1} = O(1)$ and $(\sigma^{\text{ADED}}_{S_n})^{-1} = O(1)$. Additionally, if Assumptions 3.1–3.4 hold, we have

$$\text{(iii)} \quad \frac{\sqrt{|S_n|}\left(\widehat{\text{LADE}}_{S_n}(t) - \text{LADE}_{S_n}(t)\right)}{\sigma^{\text{LADE}}_{S_n}} \xrightarrow{d} \text{Normal}(0, 1),$$

provided that

$$\frac{1}{\sigma^{\text{LADE}}_{S_n}} = O(1), \qquad \frac{\sigma^{\text{ADEY}}_{S_n}\sigma^{\text{ADED}}_{S_n}}{\sqrt{|S_n|}\sigma^{\text{LADE}}_{S_n}} = o(1), \qquad \frac{(\sigma^{\text{ADED}}_{S_n})^2}{\sqrt{|S_n|}\sigma^{\text{LADE}}_{S_n}} = o(1). \tag{4.5}$$

[5] Note that Assumption 4.7 is weaker than Assumption 3.4 of Kojevnikov *et al.* (2021). This comes from the following two facts. First, the ψ-weak dependent processes considered here are uniformly bounded by Assumptions 4.2 and 4.3, while Kojevnikov *et al.* (2021) only assume the existence of $4 + \varepsilon$ moments of them. Second, they consider a more general form of ψ-function than ours. See Assumption 2.1 of their paper and Lemma B.3.

The conditions in (4.5) are fairly mild, which are satisfied especially when the estimators are $\sqrt{|S_n|}$-consistent. Nonetheless, it should be noted that Theorem 4.2 does not rule out the possibility that the estimators exhibit slower convergence rates.

Remark 4.3 (Constant IEM)**.** Note that Assumption 4.4 does not exclude the case where the IEM is a constant function (so that $S_n = N_n$). Therefore, the results in Theorems 4.1 and 4.2 can be applied to this case as well. This means that, under the assumptions made, if we are interested only in the direct effect of own IV, we can simply ignore the others IV in the estimation. Similar results can be found in Sävje *et al.* (2021).

4.3 Statistical inference

4.3.1 Network HAC variance estimator

In this subsection, we develop the network HAC estimator and prove its asymptotic property. Under Assumption 4.5, we can see that

$$(\sigma_{S_n}^{\mathrm{ADEY}})^2 = \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\mathrm{Cov}\left[V_i^{\mathrm{ADEY}}, V_j^{\mathrm{ADEY}}\right]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant n-1\},$$

and analogous equalities hold for $(\sigma_{S_n}^{\mathrm{ADED}})^2$ and $(\sigma_{S_n}^{\mathrm{LADE}})^2$. Then, the infeasible (oracle) network HAC estimator of $(\sigma_{S_n}^{\mathrm{ADEY}})^2$ is given by

$$(\widetilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2 := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}(V_i^{\mathrm{ADEY}} - \mathbb{E}[V_i^{\mathrm{ADEY}}])(V_j^{\mathrm{ADEY}} - \mathbb{E}[V_j^{\mathrm{ADEY}}])\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\},$$

where $b_n \geqslant 0$ is a bandwidth parameter that grows as $n\to\infty$. This estimator is infeasible because both V_i^{ADEY} and $\mathbb{E}[V_i^{\mathrm{ADEY}}]$ are unobservable to us. For constructing feasible variance estimators, we compute

$$\widehat{V}_i^{\mathrm{ADEY}} := \widehat{W}_i^Y - \frac{\widehat{\mu}_{S_n}^Y(1,t)}{\widehat{p}_{S_n}(1,t)}W_i^Z + \frac{\widehat{\mu}_{S_n}^Y(0,t)}{\widehat{p}_{S_n}(0,t)}W_i^{1-Z}, \qquad \widehat{V}_i^{\mathrm{ADED}} := \widehat{W}_i^D - \frac{\widehat{\mu}_{S_n}^Y(1,t)}{\widehat{p}_{S_n}(1,t)}W_i^Z + \frac{\widehat{\mu}_{S_n}^Y(0,t)}{\widehat{p}_{S_n}(0,t)}W_i^{1-Z},$$

$$\widehat{V}_i^{\mathrm{LADE}} := \frac{1}{\widehat{\mathrm{ADED}}_{S_n}(t)}\widehat{V}_i^{\mathrm{ADEY}} - \frac{\widehat{\mathrm{ADEY}}_{S_n}(t)}{[\widehat{\mathrm{ADED}}_{S_n}(t)]^2}\widehat{V}_i^{\mathrm{ADED}},$$

where

$$\widehat{W}_i^Y := Y_i\left[\frac{\mathbf{1}\{Z_i=1,T_i=t\}}{\widehat{p}_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i=0,T_i=t\}}{\widehat{p}_{S_n}(0,t)}\right],$$
$$\widehat{W}_i^D := D_i\left[\frac{\mathbf{1}\{Z_i=1,T_i=t\}}{\widehat{p}_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i=0,T_i=t\}}{\widehat{p}_{S_n}(0,t)}\right].$$

Note that the sample mean of each of $\widehat{V}_i^{\mathrm{ADEY}}$, $\widehat{V}_i^{\mathrm{ADED}}$, and $\widehat{V}_i^{\mathrm{LADE}}$ is zero. Then, the feasible network HAC estimators are given by

$$(\widehat{\sigma}_{S_n}^{\mathrm{ADEY}})^2 := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\widehat{V}_i^{\mathrm{ADEY}}\widehat{V}_j^{\mathrm{ADEY}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\},$$

$$(\widehat{\sigma}_{S_n}^{\mathrm{ADED}})^2 := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\widehat{V}_i^{\mathrm{ADED}}\widehat{V}_j^{\mathrm{ADED}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\},$$
$$(\widehat{\sigma}_{S_n}^{\mathrm{LADE}})^2 := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\widehat{V}_i^{\mathrm{LADE}}\widehat{V}_j^{\mathrm{LADE}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\}.$$

Recall that $S_{\boldsymbol{A}}(i,s)$ denotes the subset of S_n composed of units within s distance from unit i. We write its k-th sample moment as $M_{S_n}(s,k) := |S_n|^{-1}\sum_{i\in S_n}|S_{\boldsymbol{A}}(i,s)|^k$. Further, define

$$\mathcal{J}_{S_n}(s,b_n) := \{(i,j,k,l)\in S_n^4 : \ell_{\boldsymbol{A}}(i,j)=s, \ell_{\boldsymbol{A}}(i,k)\leqslant b_n, \ell_{\boldsymbol{A}}(j,l)\leqslant b_n\}.$$

Assumption 4.8 (Weak dependence 3)**.**

(i) There exists some $0<\epsilon<1$ such that $\sum_{s=1}^{n-1}M_{S_n}^{\partial}(s)\widetilde{\theta}_{n,s}^{1-\epsilon}=O(1)$ and $\sum_{s=0}^{n-1}|\mathcal{J}_{S_n}(s,b_n)|\widetilde{\theta}_{n,s}^{1-\epsilon}=o(|S_n|^2)$.

(ii) $M_{S_n}(b_n,k)=o(|S_n|^{k/2})$ for each $k\in\{1,2\}$.

This assumption restricts both the network structure and the rate of divergence of b_n in a similar manner to Assumption 7 of Leung (2021) and Assumption 4.1 of Kojevnikov *et al.* (2021). The first part of Assumption 4.8(i) strengthens Assumption 4.6(ii) and ensures $\sqrt{|S_n|}$-consistency of our estimators (see Remark 4.2). The second part of Assumption 4.8(i) corresponds to Assumption 4.1(iii) of Kojevnikov *et al.* (2021). Assumption 4.8(ii) is the same as Assumption 7(b)–(c) of Leung (2021). Under these conditions, we can derive the probability limits of the infeasible oracle variance estimators, and evaluate the stochastic errors caused by replacing unobserved V_i^{ADEY}, V_i^{ADED}, and V_i^{LADE} with their estimators $\widehat{V}_i^{\mathrm{ADEY}}$, $\widehat{V}_i^{\mathrm{ADED}}$, and $\widehat{V}_i^{\mathrm{LADE}}$.

Let

$$B_{S_n}^{\mathrm{ADEY}} := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\mathbb{E}[V_i^{\mathrm{ADEY}}]\,\mathbb{E}[V_j^{\mathrm{ADEY}}]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\},$$
$$B_{S_n}^{\mathrm{ADED}} := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\mathbb{E}[V_i^{\mathrm{ADED}}]\,\mathbb{E}[V_j^{\mathrm{ADED}}]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\},$$
$$B_{S_n}^{\mathrm{LADE}} := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\mathbb{E}[V_i^{\mathrm{LADE}}]\,\mathbb{E}[V_j^{\mathrm{LADE}}]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\}.$$

Theorem 4.3. Suppose that Assumptions 4.1 – 4.5 and 4.8 hold. Then, if $|S_n|\to\infty$ and $b_n\to\infty$, we have

$$\text{(i)}\quad (\widehat{\sigma}_{S_n}^{\mathrm{ADEY}})^2 = (\sigma_{S_n}^{\mathrm{ADEY}})^2 + B_{S_n}^{\mathrm{ADEY}} + o_P(1),$$
$$\text{(ii)}\quad (\widehat{\sigma}_{S_n}^{\mathrm{ADED}})^2 = (\sigma_{S_n}^{\mathrm{ADED}})^2 + B_{S_n}^{\mathrm{ADED}} + o_P(1).$$

Additionally, if Assumptions 3.1–3.4 hold, we have

$$\text{(iii)}\quad (\widehat{\sigma}_{S_n}^{\mathrm{LADE}})^2 = (\sigma_{S_n}^{\mathrm{LADE}})^2 + B_{S_n}^{\mathrm{LADE}} + o_P(1).$$

In the proof, we show that

$$(\widehat{\sigma}_{S_n}^{\mathrm{ADEY}})^2 = (\widetilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2 + B_{S_n}^{\mathrm{ADEY}} + o_P(1),$$

where the infeasible oracle estimator $(\widetilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2$ is consistent for $(\sigma_{S_n}^{\mathrm{ADEY}})^2$. There is an asymptotic bias term $B_{S_n}^{\mathrm{ADEY}}$ due to the fact that we cannot estimate the heterogeneous mean $\mathbb{E}[V_i^{\mathrm{ADEY}}]$. It is well known in the design-based uncertainty framework that heterogeneous means cause standard variance estimators to have asymptotic biases (cf. Imbens and Rubin, 2015).

4.3.2 Wild bootstrap

Alternatively to the HAC estimator, we can consider using a network-dependent bootstrap method. Here, we particularly focus on Kojevnikov's (2021) wild bootstrap approach.

For exposition, we focus only on constructing a confidence interval for $\mathrm{ADEY}_{S_n}(t)$. The following procedure can be applied to the other parameters as well. As shown in (A.2) in Appendix A, we have

$$\sqrt{|S_n|}\left(\widehat{\mathrm{ADEY}}_{S_n}(t) - \mathrm{ADEY}_{S_n}(t)\right) = \frac{1}{\sqrt{|S_n|}} \sum_{i \in S_n} V_i^{\mathrm{ADEY}} + o_P(1).$$

(Note that $|S_n|^{-1/2} \sum_{i \in S_n} \mathbb{E}[V_i^{\mathrm{ADEY}}] = 0$.) Thus, if it is possible to simulate the distribution of $|S_n|^{-1/2} \sum_{i \in S_n} V_i^{\mathrm{ADEY}}$, we can construct an asymptotically valid confidence interval for $\mathrm{ADEY}_{S_n}(t)$. To this end, noting that the sample mean of $\widehat{V}_i^{\mathrm{ADEY}}$ over S_n is zero, we construct a bootstrap counterpart $V_i^{*,\mathrm{ADEY}}$ of V_i^{ADEY} in the following procedure: $V_i^{*,\mathrm{ADEY}} := \widehat{V}_i^{\mathrm{ADEY}} R_i$, where R_i is the i-th element of the $|S_n| \times 1$ vector $[\Omega_{S_n}(b_n)]^{1/2} \zeta_{S_n}$ with

$$\Omega_{S_n}(b_n) := \left(\frac{|S_{\boldsymbol{A}}(i, b_n) \cap S_{\boldsymbol{A}}(j, b_n)|}{M_{S_n}(b_n, 1)}\right)_{i,j \in S_n},$$

ζ_{S_n} is an $|S_n| \times 1$ vector of random variables drawn from $\mathrm{Normal}(0, I_{|S_n|})$ independently of the data, and b_n is a bandwidth parameter. Then, by repeatedly drawing ζ_{S_n} many times, we can obtain the distribution of $|S_n|^{-1/2} \sum_{i \in S_n} V_i^{*,\mathrm{ADEY}}$ conditional on the observed data, which serves as an approximation of the distribution of $|S_n|^{-1/2} \sum_{i \in S_n} V_i^{\mathrm{ADEY}}$. An intuition for the (first-order) validity of this bootstrap method is as follows. Since the conditional expectation of $V_i^{*,\mathrm{ADEY}}$ given the observed data is zero, we have

$$\mathrm{Var}\left[\frac{1}{\sqrt{|S_n|}} \sum_{i \in S_n} V_i^{*,\mathrm{ADEY}} \,\middle|\, \text{data}\right] = \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{j \in S_n} \widehat{V}_i^{\mathrm{ADEY}} \widehat{V}_j^{\mathrm{ADEY}} [\Omega_{S_n}(b_n)]_{i,j}.$$

Thus, this is a version of the HAC estimator with kernel $\Omega_{S_n}(b_n)$. For more details, see Kojevnikov (2021).

5 Numerical Illustrations

5.1 Monte Carlo simulation

In this subsection, we investigate the finite sample properties of our methods using a set of Monte Carlo experiments. The following two DGPs are considered:

$$\textbf{DGP 1:}\quad Y_i = \beta_{0i} + \beta_{1i} D_i$$
$$D_i = \mathbf{1}\left\{\gamma_{0i} + \gamma_1 Z_i + \gamma_2 \sum_{j \neq i} A_{ij}^{(K)} Z_j \geqslant 0\right\}$$

where β_{0i}, β_{1i}, and γ_{0i} are drawn from $\text{Normal}(1,1)$, $\text{Normal}(1,1)$, and $\text{Normal}(-2,1)$, respectively, $(\gamma_1, \gamma_2) = (1.5, 0.5)$, and Z_i's are IID $\text{Bernoulli}(0.4)$.

$$\textbf{DGP 2:}\quad Y_i = \beta_{0i} + \beta_{1i} D_i + \beta_2 \sum_{j \neq i} A_{ij}^{(K)} D_j$$
$$D_i = \mathbf{1}\left\{\gamma_{0i} + \gamma_1 Z_i \geqslant 0\right\}$$

where β_{0i}, β_{1i}, and γ_{0i} are drawn from $\text{Normal}(1,1)$, $\text{Normal}(1,1)$, and $\text{Normal}(-1.5,1)$, respectively, $(\beta_2, \gamma_1) = (0.5, 1.5)$, and Z_i's are IID $\text{Bernoulli}(0.4)$. The individual-specific coefficients are drawn only once, and they are fixed throughout the simulations. Here, $A_{ij}^{(K)}$ is a ring-shape network where individuals interact with their K-nearest neighbors:

$$A_{ij}^{(K)} = \begin{cases} 1 & \text{if } \min\{|i-j|, |i-j+n|, |i-j-n|\} \leqslant K \\ 0 & \text{otherwise} \end{cases}$$

For both DGPs, we consider two cases $K \in \{2, 3\}$. For the specification of the IEM, we consider two versions for each DGP in which the one is a correctly specified IEM and the other is misspeficied:

DGP 1:	Correct IEM	$T_i = \sum_{j \neq i} A_{ij}^{(K)} Z_j$
	Incorrect IEM	$T_i = \sum_{j \neq i} A_{ij}^{(1)} Z_j$
DGP 2:	Correct IEM	$T_i = \sum_{j \neq i} A_{ij}^{(K)} D_j = \sum_{j \neq i} A_{ij}^{(K)} \mathbf{1}\{\gamma_{0j} + \gamma_1 Z_j \geqslant 0\}$
	Incorrect IEM	$T_i = \sum_{j \neq i} A_{ij}^{(1)} D_j = \sum_{j \neq i} A_{ij}^{(1)} \mathbf{1}\{\gamma_{0j} + \gamma_1 Z_j \geqslant 0\}$

The forms of $\mu_i^D(z,t)$ and $\mu_i^Y(z,t)$ under correct IEMs are straightforward. For DGP 1, when the IEM is misspecified, noting that $\sum_{j \neq i} A_{ij}^{(K)} Z_j = T_i + \sum_{j \neq i}(A_{ij}^{(K)} - A_{ij}^{(1)}) Z_j$ and the second term on the right-hand side is distributed as $\text{Binomial}(2K-2, 0.4)$, we have

$$\mu_i^Y(z,t) = \beta_{0i} + \beta_{1i} \Pr\left(\widetilde{Z}^{(K)} \geqslant -(\gamma_{0i} + \gamma_1 z + \gamma_2 t)/\gamma_2\right),$$

where $\widetilde{Z}^{(K)} \sim \text{Binomial}(2K-2, 0.4)$. Similarly, for DGP 2, we have under the incorrect IEM that

$$\mu_i^Y(z,t) = \beta_{0i} + \beta_1 \mathbf{1}\{\gamma_{0i} + \gamma_1 z \geqslant 0\} + \beta_2 t + \beta_2 \sum_{j \neq i} (A_{ij}^{(K)} - A_{ij}^{(1)}) \Pr(Z \geqslant -\gamma_{0j}/\gamma_1),$$

where $\Pr(Z \geqslant -\gamma_{0j}/\gamma_1) = \mathbf{1}\{-\gamma_{0j}/\gamma_1 \leqslant 0\} + 0.4 \cdot \mathbf{1}\{0 < -\gamma_{0j}/\gamma_1 \leqslant 1\}$.

The data are generated for two sample sizes $n \in \{500, 1000\}$. Note that in our DGPs, all individuals have the same network structure and the same distribution of T_i. Thus, we use the whole sample N_n as S_n. For each setup, we estimate $(\text{ADEY}_{N_n}(2), \text{LADE}_{N_n}(2))$ in DGP 1 and $(\text{ADEY}_{N_n}(1), \text{LADE}_{N_n}(1))$ in DGP 2 using the estimators introduced in Section 4. The performance of the estimators is measured in terms of the bias and the root mean squared error (RMSE) based on 1,000 Monte Carlo repetitions.

The results are summarized in Table 1. We can find that our estimators work satisfactorily well overall irrespective of whether the IEM is correctly- or mis-specified. Although our estimators are not unbiased for the finite sample as we have stated in Remark 4.1, the biases are sufficiently small in all setups. The RMSE values for the LADE parameter are larger than that of ADEY. This is because the estimation of LADE involves the estimation of ADED (i.e., the average probability of compliance conditional on $T_i = t$). Note that the estimation accuracy of ADED and the size of the compliers depend largely on the specification of the IEM T and the choice of value t. Depending on these factors, especially for small n, ADED may be estimated to be zero, resulting in the failure of LADE estimation.

Table 1: Bias and RMSE

			Correct IEM				Incorrect IEM			
			ADEY		LADE		ADEY		LADE	
DGP	K	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
1	2	500	-0.0008	0.1995	-0.0034	0.3410	-0.0103	0.3184	-0.0252	0.6099
		1000	0.0021	0.1456	-0.0005	0.2554	0.0046	0.2165	0.0054	0.3892
	3	500	-0.0034	0.2002	-0.0080	0.3460	-0.0118	0.3212	-0.0420	0.8275
		1000	-0.0044	0.1511	-0.0107	0.2694	0.0028	0.2227	0.0076	0.4697
2	2	500	0.0055	0.1699	0.0071	0.3528	0.0001	0.1672	-0.0052	0.3449
		1000	-0.0065	0.1212	-0.0221	0.2754	-0.0008	0.1208	-0.0031	0.2709
	3	500	0.0106	0.1785	0.0307	0.3742	-0.0010	0.1708	-0.0078	0.3524
		1000	0.0059	0.1378	-0.0045	0.3000	-0.0021	0.1230	-0.0060	0.2768

We next examine the performance of the HAC estimator and the wild bootstrap approach introduced in subsections 4.3.1 and 4.3.2, respectively. The DGPs and the target parameters considered are the same as above. For the ADEY and LADE parameters in each setup, we compute the coverage rate of the 95% confidence interval obtained based on these two approaches. The bandwidth is chosen from $b_n \in \{K, 2K\}$ for both approaches.

The results are summarized in Table 2. Overall, we can see that the empirical coverage ratios are reasonably close to the nominal 95% level for both estimators. Given that the estimators have non-negligible biases, the above results indicate that the magnitude of the bias is not severe. For some specific designs and parameters (such as ADEY in DGP 1 under the incorrect IEM), the confidence intervals tend to be narrower than the nominal level, but it seems possible to correct the size distortion by increasing the sample size.

Table 2: Coverage ratio of the 95% CI

				Correct IEM				Incorrect IEM			
				ADEY		LADE		ADEY		LADE	
DGP	K	n	b_n	HAC	Bootstrap	HAC	Bootstrap	HAC	Bootstrap	HAC	Bootstrap
1	2	500	2	0.939	0.935	0.943	0.939	0.922	0.921	0.950	0.950
			4	0.936	0.926	0.931	0.932	0.909	0.916	0.945	0.946
		1000	2	0.945	0.946	0.954	0.950	0.952	0.951	0.952	0.948
			4	0.941	0.942	0.944	0.950	0.944	0.942	0.945	0.943
	3	500	3	0.935	0.938	0.942	0.945	0.916	0.912	0.940	0.939
			6	0.915	0.919	0.934	0.933	0.898	0.902	0.918	0.926
		1000	3	0.945	0.936	0.943	0.932	0.934	0.934	0.950	0.951
			6	0.932	0.925	0.934	0.922	0.937	0.931	0.945	0.943
2	2	500	2	0.954	0.955	0.951	0.949	0.947	0.940	0.946	0.945
			4	0.951	0.953	0.947	0.941	0.931	0.937	0.936	0.933
		1000	2	0.953	0.949	0.952	0.951	0.954	0.948	0.956	0.951
			4	0.949	0.949	0.950	0.946	0.946	0.943	0.951	0.949
	3	500	3	0.938	0.935	0.940	0.933	0.946	0.946	0.942	0.941
			6	0.918	0.926	0.912	0.920	0.933	0.937	0.930	0.939
		1000	3	0.929	0.933	0.935	0.936	0.946	0.942	0.955	0.959
			6	0.929	0.924	0.932	0.932	0.942	0.940	0.956	0.944

5.2 Empirical application

We apply the proposed methods to the data from Paluck *et al.*'s (2016) field experiment on anti-conflict intervention programs at American middle schools. During the 2012-2013 school year, the research team organized intervention meetings to help students identify common conflict behaviors in their schools and instruct them on behavioral strategies to mitigate conflicts. The purpose of the experiment was to examine how the intervention program affects participants' behavior and whether the students' social networks influence the climate of conflict in schools.

The data include $n = 24,471$ students in 56 public middle schools in the state of New Jersey. Half of these schools were randomly selected to host the anti-conflict intervention program. Within each selected school, a group of students (called *seed-eligible students*) were non-randomly selected by the research team, and half of these students (called *seed students* or *treatment-eligible students*) were randomly invited to join the program. The experimental design was one-sided noncompliance where meeting attendance was not compulsory and the students without an invitation were not able to attend. Thus, there are only compliers and never-takers in this empirical analysis, and joining in the intervention program means that the student is a complier.

Before starting the intervention program, the research team measured the students' social networks by asking them to nominate up to 10 students in their school with whom they had spent time in person or online in the past few weeks. We construct a symmetric adjacency matrix $\boldsymbol{A}$ by treating the pair of students as friends if either student nominated the other, as in Aronow and Samii (2017).

In our analysis, $Z_i \in \{0, 1\}$ indicates whether student i received an invitation to the intervention program (i.e., whether student i was a seed student), and $D_i \in \{0, 1\}$ represents the participation in the intervention program (i.e., whether student i attended at least one intervention meeting). Let $Y_i \in \{0, 1\}$ be a (self-reported) indicator

Table 3: Descriptive statistics for seed-eligible students

	Overall	Treatment eligibility $Z_i = 0$	$Z_i = 1$
Treatment			
$D_i = 0$	2,357 (79%)	1,491 (100%)	866 (58%)
$D_i = 1$	626 (21%)	0 (0%)	626 (42%)
Outcome			
$Y_i = 0$	2,684 (90%)	1,392 (93%)	1,292 (87%)
$Y_i = 1$	299 (10%)	99 (7%)	200 (13%)
IEM1			
$T_{1i} = 0$	1,435 (48%)	716 (48%)	719 (48%)
$T_{1i} = 1$	1,548 (52%)	775 (52%)	773 (52%)
IEM2			
$T_{2i} = 0$	2,277 (76%)	1,144 (77%)	1,133 (76%)
$T_{2i} = 1$	706 (24%)	347 (23%)	359 (24%)

for the wearing of a program wristband given by the treated students as a reward to students for engaging in friendly or conflict-mitigating behaviors. This is regarded as a proxy variable of student's willingness to endorse anti-conflict norms and behaviors, and the same outcome variable is used in Aronow and Samii (2017) and Leung (2021). We consider the following two IEMs: $T_{1i} = \mathbf{1}\{\sum_{j \neq i} A_{ij} Z_j > 0\}$ and $T_{2i} = \mathbf{1}\{\sum_{j \neq i} A_{ij} D_j > 0\}$, respectively labeled as "IEM1" and "IEM2". In other words, T_{1i} and T_{2i} indicate whether student i has at least one treated and treatment-eligible friend, respectively. In line with Assumption 4.1(ii), we focus on the following sub-populations:

$$S_n(\delta) = \{i \in N_n : i \text{ is a seed-eligible student who has } \delta \text{ seed-eligible friend(s)}\} \quad \text{for } \delta \in \{1, 2, 3\}.$$

The descriptive statistics for the seed-eligible students are summarized in Table 3. We can see that all students without invitation did not actually join the intervention program, implying that the monotonicity condition in Assumption 3.3 holds. It is also interesting that not just IEM1 but the distribution of IEM2 is also insensitive to the student's own invitation status. This would suggest that one's treatment eligibility does not have substantial impacts on the others' treatment choices. Indeed, we found that the conditional distribution of $\sum_{j \neq i} A_{ij} D_j$ given $Z_i = 1$ is almost identical to that given $Z_i = 0$.

For each realization of both IEMs, Tables 4 and 5 present the ITT estimates with the standard errors based on the network HAC estimation using bandwidth $b_n \in \{0, 1, 2, 3\}$.[6] Overall, receiving an invitation has a statistically significant positive effect on the probability of wearing a wristband, which is consistent with previous findings (e.g., Aronow and Samii, 2017; Leung, 2021). For example, the estimate of $\mathrm{ADEY}_{S_n(1)}(0)$ for IEM1 indicates that receiving an invitation leads to about a six percentage point increase in the probability of wearing a wristband for the seed-eligible students whose seed-eligible friend is not treatment-eligible. Similarly, the ADED estimates indicate positive effects of receiving an invitation on the probability of participation, which

[6] The wild bootstrap produced similar standard errors to those reported here. To save space, we omit the results from the wild bootstrap.

supports the IV relevance condition in Assumption 3.2. Remarkably, $\mathrm{ADEY}_{S_n(\delta)}(1)$ is substantially larger than $\mathrm{ADEY}_{S_n(\delta)}(0)$ for IEM2, implying that having at least one treated friend boosts the direct effect. This is reasonable because the wristband is given by the program participants.

The estimates of LADE and their standard errors based on the HAC estimation are also reported in Tables 4 and 5. For example, the estimate of $\mathrm{LADE}_{S_n(1)}(1)$ based on IEM1 indicates a twenty-four percentage point increase in the probability of wearing a wristband for the seed-eligible students who have a treatment-eligible friend. Interestingly, the LADE estimates tend to be larger than the corresponding ITT estimates, implying that the ITT analysis might underestimate the effect of the anti-conflict intervention program.

Nonetheless, we should be cautious in interpreting the LADE estimates because the interpretation of LADE crucially depends on which sufficient condition we consider for Assumption 3.4. Due to the nature of the anti-conflict intervention program, it is plausible to imagine that the never-takers (i.e., those who never join the intervention program irrespective of their invitation status) were unable to affect the participation of others. Thus, the third sufficient condition (3.3) for Assumption 3.4 is plausible here, and LADE aggregates the ADE of participation in the intervention program and the ASE caused by changing the student's own treatment eligibility. However, since one's treatment eligibility seems to have little impact on the others' treatment choice as observed above, LADE should mainly account for the ADE of the intervention program.

6 Conclusion

In this study, we developed a causal inference method that simultaneously addresses cross-unit interference within a social network and the issue of noncompliance with the assigned treatment. The key feature of our approach is to admit the possibility of misspecification of IEM, which is a function of IVs that enables summarizing the spillover effects into a low-dimensional variable. We conducted the identification analysis for the ITT effect and the ADE for compliers. Here, we mainly discussed a causal interpretation for the ADE parameter and the identification result for the LADE parameter. Based on the identification results, we proposed nonparametric procedures for estimating the treatment parameters and investigated their asymptotic properties based on the ANI framework originally introduced by Leung (2021). We also considered the statistical inference methods based on the network HAC estimation and the wild bootstrap. The empirical application to the data of Paluck *et al.* (2016) highlighted the usefulness of our method.

Several important research topics related to our study remain to be investigated. First, it would be of interest to examine whether some treatment parameters can be recovered in the case where some of our identification conditions are violated. In that case, it would be difficult to achieve point identification of the treatment parameters, and a promising approach in this direction would be to pursue a partial identification strategy (cf. Manski, 2013). Second, we could extend our analysis to the situation in which the treatment and/or IV take non-binary values. It is known in the absence of interference that the standard monotonicity condition does not apply to a case with a discrete treatment or IV, and a careful analysis would be required for handling this issue (cf. Heckman and Pinto, 2018). Finally, our asymptotic theory depends somewhat on the sparsity of the network, and it may be worthwhile to investigate under what conditions it is possible (or impossible) to derive similar results for dense networks.

Table 4: Direct effects conditional on $T_{1i} = 0$ or $T_{2i} = 0$

S_n	$\lvert S_n \rvert$	b_n	$\text{ADEY}_{S_n}(0)$ Estimate	$\text{ADEY}_{S_n}(0)$ SE	$\text{ADED}_{S_n}(0)$ Estimate	$\text{ADED}_{S_n}(0)$ SE	$\text{LADE}_{S_n}(0)$ Estimate	$\text{LADE}_{S_n}(0)$ SE
(A) IEM1								
$S_n(1)$	1006	0	0.063	0.027	0.423	0.031	0.150	0.060
		1		0.027		0.031		0.060
		2		0.028		0.041		0.064
		3		0.030		0.062		0.063
$S_n(2)$	660	0	0.101	0.044	0.365	0.052	0.277	0.108
		1		0.044		0.052		0.109
		2		0.039		0.065		0.092
		3		0.045		0.076		0.104
$S_n(3)$	341	0	−0.034	0.103	0.500	0.107	−0.068	0.209
		1		0.104		0.107		0.211
		2		0.106		0.116		0.215
		3		0.106		0.124		0.215
(B) IEM2								
$S_n(1)$	1006	0	0.039	0.019	0.299	0.023	0.131	0.060
		1		0.019		0.024		0.060
		2		0.020		0.034		0.062
		3		0.022		0.054		0.067
$S_n(2)$	660	0	0.043	0.020	0.198	0.028	0.219	0.093
		1		0.021		0.028		0.093
		2		0.019		0.035		0.085
		3		0.020		0.044		0.088
$S_n(3)$	341	0	−0.012	0.025	0.130	0.032	−0.094	0.202
		1		0.025		0.033		0.202
		2		0.025		0.037		0.202
		3		0.022		0.040		0.172

Acknowledgments

The authors thank the participants of seminars at Kobe University and Osaka University for their beneficial comments. This work was supported by JSPS KAKENHI Grant Numbers 19H01473 and 20K01597. The data set used in this study is available through the Inter-university Consortium for Political and Social Research (Paluck *et al.*, 2020).

Table 5: Direct effects conditional on $T_{1i} = 1$ or $T_{2i} = 1$

			$\mathrm{ADEY}_{S_n}(1)$		$\mathrm{ADED}_{S_n}(1)$		$\mathrm{LADE}_{S_n}(1)$	
S_n	$\lvert S_n \rvert$	b_n	Estimate	SE	Estimate	SE	Estimate	SE
(A) IEM1								
$S_n(1)$	1006	0	0.102	0.028	0.423	0.031	0.240	0.062
		1		0.029		0.034		0.061
		2		0.028		0.045		0.060
		3		0.028		0.061		0.052
$S_n(2)$	660	0	0.098	0.029	0.496	0.032	0.198	0.056
		1		0.030		0.035		0.056
		2		0.035		0.047		0.064
		3		0.033		0.063		0.059
$S_n(3)$	341	0	0.022	0.034	0.400	0.040	0.054	0.083
		1		0.035		0.047		0.085
		2		0.038		0.057		0.093
		3		0.037		0.070		0.091
(B) IEM2								
$S_n(1)$	1006	0	0.235	0.054	0.864	0.033	0.273	0.062
		1		0.054		0.033		0.061
		2		0.056		0.035		0.065
		3		0.052		0.044		0.058
$S_n(2)$	660	0	0.185	0.052	0.897	0.027	0.206	0.058
		1		0.052		0.027		0.057
		2		0.058		0.028		0.064
		3		0.052		0.028		0.057
$S_n(3)$	341	0	0.075	0.069	0.891	0.039	0.084	0.077
		1		0.071		0.039		0.079
		2		0.081		0.038		0.091
		3		0.081		0.041		0.091

A Appendix: Proofs

A.1 Proof of Proposition 3.1

We prove only the results for the outcome variable, and those for the treatment receipt can be shown in the same manner. We first note that the observed outcome can be written as

$$Y_i = \sum_{z_i=0}^{1} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{Z_i = z_i, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}\} y_i(z_i, \boldsymbol{z}_{-i}).$$

We then observe that

$$\begin{aligned} \mu_i^Y(z,t) &= \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} y_i(z, \boldsymbol{z}_{-i}) \Pr[\boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i} \mid Z_i = z, T_i = t] \\ &= \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} y_i(z, \boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t), \end{aligned} \tag{A.1}$$

where the second line follows from Assumption 3.1. This equality implies the results for $\mathrm{ADEY}_{S_n}(t)$ and $\mathrm{ASEY}_{S_n}(z,t,t')$. □

A.2 Proof of Theorem 3.1

Observe that $D_i = \sum_{z_i=0}^{1} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{Z_i = z_i, \boldsymbol{Z}_{-i} = \boldsymbol{z}_{-i}\} D_i(z_i, \boldsymbol{z}_{-i})$. By Assumption 3.1, it holds that

$$\begin{aligned} \mu_i^D(z,t) &= \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} D_i(z, \boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t) \\ &= \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) \neq D_i(0, \boldsymbol{z}_{-i})\} D_i(z, \boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t) \\ &\quad + \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i})\} D_i(z, \boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t). \end{aligned}$$

Thus, Assumption 3.3 implies that

$$\begin{aligned} \mathrm{ADED}_{S_n}(t) &= \frac{1}{|S_n|} \sum_{i \in S_n} [\mu_i^D(1,t) - \mu_i^D(0,t)] \\ &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) \neq D_i(0, \boldsymbol{z}_{-i})\}\{D_i(1, \boldsymbol{z}_{-i}) - D_i(0, \boldsymbol{z}_{-i})\} \pi_i(\boldsymbol{z}_{-i}, t) \\ &\quad + \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathbf{1}\{D_i(1, \boldsymbol{z}_{-i}) = D_i(0, \boldsymbol{z}_{-i})\}\{D_i(1, \boldsymbol{z}_{-i}) - D_i(0, \boldsymbol{z}_{-i})\} \pi_i(\boldsymbol{z}_{-i}, t) \\ &= \frac{1}{|S_n|} \sum_{i \in S_n} \sum_{\boldsymbol{z}_{-i} \in \{0,1\}^{n-1}} \mathcal{C}_i(\boldsymbol{z}_{-i}) \pi_i(\boldsymbol{z}_{-i}, t). \end{aligned}$$

In the same manner, we can show that

$$\begin{aligned}
\mathrm{ADEY}_{S_n}(t) &= \frac{1}{|S_n|}\sum_{i\in S_n}[\mu_i^Y(1,t)-\mu_i^Y(0,t)] \\
&= \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{\boldsymbol{z}_{-i}\in\{0,1\}^{n-1}}\{y_i(1,\boldsymbol{z}_{-i})-y_i(0,\boldsymbol{z}_{-i})\}\mathbf{1}\{D_i(1,\boldsymbol{z}_{-i})\neq D_i(0,\boldsymbol{z}_{-i})\}\pi_i(\boldsymbol{z}_{-i},t) \\
&\quad + \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{\boldsymbol{z}_{-i}\in\{0,1\}^{n-1}}\{y_i(1,\boldsymbol{z}_{-i})-y_i(0,\boldsymbol{z}_{-i})\}\mathbf{1}\{D_i(1,\boldsymbol{z}_{-i})= D_i(0,\boldsymbol{z}_{-i})\}\pi_i(\boldsymbol{z}_{-i},t) \\
&= \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{\boldsymbol{z}_{-i}\in\{0,1\}^{n-1}}\{y_i(1,\boldsymbol{z}_{-i})-y_i(0,\boldsymbol{z}_{-i})\}\mathcal{C}_i(\boldsymbol{z}_{-i})\pi_i(\boldsymbol{z}_{-i},t),
\end{aligned}$$

where the last line follows from Assumptions 3.3 and 3.4. Combining these equalities with Assumption 3.2, we obtain the desired result. □

A.3 Proof of Theorem 4.2

Proof of result (i). Observe that

$$\begin{aligned}
\widehat{\mu}^Y_{S_n}(z,t) &= \frac{\breve{\mu}^Y_{S_n}(z,t)}{\widehat{p}_{S_n}(z,t)}p_{S_n}(z,t) \\
&= \breve{\mu}^Y_{S_n}(z,t) - \frac{\breve{\mu}^Y_{S_n}(z,t)}{\widehat{p}_{S_n}(z,t)}[\widehat{p}_{S_n}(z,t)-p_{S_n}(z,t)] \\
&= \breve{\mu}^Y_{S_n}(z,t) - \frac{\bar{\mu}^Y_{S_n}(z,t)}{p_{S_n}(z,t)}[\widehat{p}_{S_n}(z,t)-p_{S_n}(z,t)] - \frac{\breve{\mu}^Y_{S_n}(z,t)-\bar{\mu}^Y_{S_n}(z,t)}{\widehat{p}_{S_n}(z,t)}[\widehat{p}_{S_n}(z,t)-p_{S_n}(z,t)] \\
&\quad + \bar{\mu}^Y_{S_n}(z,t)\left(\frac{1}{p_{S_n}(z,t)}-\frac{1}{\widehat{p}_{S_n}(z,t)}\right)[\widehat{p}_{S_n}(z,t)-p_{S_n}(z,t)] \\
&= \breve{\mu}^Y_{S_n}(z,t) - \frac{\bar{\mu}^Y_{S_n}(z,t)}{p_{S_n}(z,t)}[\widehat{p}_{S_n}(z,t)-p_{S_n}(z,t)] + o_P\left(\frac{1}{\sqrt{|S_n|}}\right),
\end{aligned}$$

where the last equality holds from Lemmas B.1 and B.2 and Assumption 4.3. Using this, we have

$$\begin{aligned}
&\widehat{\mathrm{ADEY}}_{S_n}(t) - \mathrm{ADEY}_{S_n}(t) \\
&= \widehat{\mu}^Y_{S_n}(1,t) - \widehat{\mu}^Y_{S_n}(0,t) - \bar{\mu}^Y_{S_n}(1,t) + \bar{\mu}^Y_{S_n}(0,t) \\
&= \breve{\mu}^Y_{S_n}(1,t) - \breve{\mu}^Y_{S_n}(0,t) - \frac{\bar{\mu}^Y_{S_n}(1,t)}{p_{S_n}(1,t)}[\widehat{p}_{S_n}(1,t)-p_{S_n}(1,t)] + \frac{\bar{\mu}^Y_{S_n}(0,t)}{p_{S_n}(0,t)}[\widehat{p}_{S_n}(0,t)-p_{S_n}(0,t)] \\
&\quad - \bar{\mu}^Y_{S_n}(1,t) + \bar{\mu}^Y_{S_n}(0,t) + o_P\left(\frac{1}{\sqrt{|S_n|}}\right) \\
&= \frac{1}{|S_n|}\sum_{i\in S_n}\left([W_i^Y-\mathbb{E}\,W_i^Y] - \frac{\bar{\mu}^Y_{S_n}(1,t)}{p_{S_n}(1,t)}[W_i^Z-\mathbb{E}\,W_i^Z] + \frac{\bar{\mu}^Y_{S_n}(0,t)}{p_{S_n}(0,t)}[W_i^{1-Z}-\mathbb{E}\,W_i^{1-Z}]\right) + o_P\left(\frac{1}{\sqrt{|S_n|}}\right) \\
&= \frac{1}{|S_n|}\sum_{i\in S_n}\left(V_i^{\mathrm{ADEY}} - \mathbb{E}[V_i^{\mathrm{ADEY}}]\right) + o_P\left(\frac{1}{\sqrt{|S_n|}}\right).
\end{aligned} \tag{A.2}$$

By Lemma B.4, $\{V_i^{\mathrm{ADEY}}\}_{i \in S_n}$ is ψ-weakly dependent with the dependence coefficients $\{\widetilde{\theta}_{n,s}\}_{s \geqslant 0}$. Then, letting $\widetilde{G}_{S_n}^{\mathrm{ADEY}} := |S_n|^{-1/2} \sum_{i \in S_n} (V_i^{\mathrm{ADEY}} - \mathbb{E}[V_i^{\mathrm{ADEY}}]) / \sigma_{S_n}^{\mathrm{ADEY}}$, the same arguments as in the proofs of Lemmas A.2 and A.3 of Kojevnikov *et al.* (2021) show that there exists a positive constant $C > 0$ such that

$$\sup_{a \in \mathbb{R}} \left| \Pr\left(\widetilde{G}_{S_n}^{\mathrm{ADEY}} \leqslant a \right) - \Phi(a) \right| \leqslant C \sum_{k=1}^{2} \left(\sqrt{ \frac{1}{|S_n|^{k/2} (\sigma_{S_n}^{\mathrm{ADEY}})^{2+k}} \sum_{s=0}^{n-1} c_{S_n}(s, m_n; k) \widetilde{\theta}_{n,s}^{1-\varepsilon} } + \frac{|S_n|^{k/2}}{(\sigma_{S_n}^{\mathrm{ADEY}})^k} \widetilde{\theta}_{n,m_n}^{1-\varepsilon} \right),$$

where Φ denotes the cumulative distribution function of $\mathrm{Normal}(0,1)$, and m_n and ε are as given in Assumption 4.7. The right-hand side converges to zero by Assumption 4.7, implying that $\widetilde{G}_{S_n}^{\mathrm{ADEY}} \xrightarrow{d} \mathrm{Normal}(0,1)$. Thus, we have

$$\frac{\sqrt{|S_n|} \left(\widehat{\mathrm{ADEY}}_{S_n}(t) - \mathrm{ADEY}_{S_n}(t) \right)}{\sigma_{S_n}^{\mathrm{ADEY}}} = \widetilde{G}_{S_n}^{\mathrm{ADEY}} + o_P\left(\frac{1}{\sigma_{S_n}^{\mathrm{ADEY}}} \right) \xrightarrow{d} \mathrm{Normal}(0,1),$$

under the condition $(\sigma_{S_n}^{\mathrm{ADEY}})^{-1} = O(1)$.

Proof of result (ii). Result (ii) can be shown in the same manner as in result (i).

Proof of result (iii). We can observe that

$$\begin{aligned}
&\widehat{\mathrm{LADE}}_{S_n}(t) - \mathrm{LADE}_{S_n}(t) \\
&= \frac{1}{\mathrm{ADED}_{S_n}(t)} [\widehat{\mathrm{ADEY}}_{S_n}(t) - \mathrm{ADEY}_{S_n}(t)] - \frac{\widehat{\mathrm{ADEY}}_{S_n}(t)}{\widehat{\mathrm{ADED}}_{S_n}(t) \mathrm{ADED}_{S_n}(t)} [\widehat{\mathrm{ADED}}_{S_n}(t) - \mathrm{ADED}_{S_n}(t)] \\
&= \frac{1}{\mathrm{ADED}_{S_n}(t)} [\widehat{\mathrm{ADEY}}_{S_n}(t) - \mathrm{ADEY}_{S_n}(t)] - \frac{\mathrm{ADEY}_{S_n}(t)}{[\mathrm{ADED}_{S_n}(t)]^2} [\widehat{\mathrm{ADED}}_{S_n}(t) - \mathrm{ADED}_{S_n}(t)] \\
&\quad - \left(\frac{\widehat{\mathrm{ADEY}}_{S_n}(t)}{\widehat{\mathrm{ADED}}_{S_n}(t) \mathrm{ADED}_{S_n}(t)} - \frac{\mathrm{ADEY}_{S_n}(t)}{[\mathrm{ADED}_{S_n}(t)]^2} \right) [\widehat{\mathrm{ADED}}_{S_n}(t) - \mathrm{ADED}_{S_n}(t)] \\
&= \frac{1}{|S_n|} \sum_{i \in S_n} \left(\frac{1}{\mathrm{ADED}_{S_n}(t)} [V_i^{\mathrm{ADEY}} - \mathbb{E}\, V_i^{\mathrm{ADEY}}] - \frac{\mathrm{ADEY}_{S_n}(t)}{[\mathrm{ADED}_{S_n}(t)]^2} [V_i^{\mathrm{ADED}} - \mathbb{E}\, V_i^{\mathrm{ADED}}] \right) \\
&\quad + O_P\left(\frac{\sigma_{S_n}^{\mathrm{ADEY}} \sigma_{S_n}^{\mathrm{ADED}}}{|S_n|} \right) + O_P\left(\frac{(\sigma_{S_n}^{\mathrm{ADED}})^2}{|S_n|} \right) + o_P\left(\frac{1}{\sqrt{|S_n|}} \right) \\
&= \frac{1}{|S_n|} \sum_{i \in S_n} \left(V_i^{\mathrm{LADE}} - \mathbb{E}[V_i^{\mathrm{LADE}}] \right) + O_P\left(\frac{\sigma_{S_n}^{\mathrm{ADEY}} \sigma_{S_n}^{\mathrm{ADED}}}{|S_n|} \right) + O_P\left(\frac{(\sigma_{S_n}^{\mathrm{ADED}})^2}{|S_n|} \right) + o_P\left(\frac{1}{\sqrt{|S_n|}} \right),
\end{aligned}$$

where the third line follows from Assumption 3.2 and results (i)–(ii). Here, V_i^{LADE} is uniformly bounded by Assumptions 3.2, 4.2, and 4.3, and Lemma B.4 implies that $\{V_i^{\mathrm{LADE}}\}_{i \in S_n}$ is ψ-weakly dependent with the dependence coefficients $\{\widetilde{\theta}_{n,s}\}_{s \geqslant 0}$. Then, letting $\widetilde{G}_{S_n}^{\mathrm{LADE}} := |S_n|^{-1/2} \sum_{i \in S_n} (V_i^{\mathrm{LADE}} - \mathbb{E}[V_i^{\mathrm{LADE}}]) / \sigma_{S_n}^{\mathrm{LADE}}$, the same arguments as in the proof of result (i) show that $\widetilde{G}_{S_n}^{\mathrm{LADE}} \xrightarrow{d} \mathrm{Normal}(0,1)$. Thus, in conjunction with

(4.5), we obtain

$$\frac{\sqrt{|S_n|}\left(\widehat{\mathrm{LADE}}_{S_n}(t) - \mathrm{LADE}_{S_n}(t)\right)}{\sigma_{S_n}^{\mathrm{LADE}}} = \tilde{G}_{S_n}^{\mathrm{LADE}} + O_P\left(\frac{\sigma_{S_n}^{\mathrm{ADEY}}\sigma_{S_n}^{\mathrm{ADED}}}{\sqrt{|S_n|}\sigma_{S_n}^{\mathrm{LADE}}}\right) + O_P\left(\frac{(\sigma_{S_n}^{\mathrm{ADED}})^2}{\sqrt{|S_n|}\sigma_{S_n}^{\mathrm{LADE}}}\right) + o_P\left(\frac{1}{\sigma_{S_n}^{\mathrm{LADE}}}\right)$$
$$\xrightarrow{d} \mathrm{Normal}(0,1).$$

□

A.4 Proof of Theorem 4.3

To save space, we prove only the result for $\sigma_{S_n}^{\mathrm{ADEY}}$ (those for $\sigma_{S_n}^{\mathrm{ADED}}$ and $\sigma_{S_n}^{\mathrm{LADE}}$ can be shown in the same manner). It is easy to see that

$$\begin{aligned}(\hat{\sigma}_{S_n}^{\mathrm{ADEY}})^2 &= (\tilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2 + B_{S_n}^{\mathrm{ADEY}} \\ &\quad + \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(\hat{V}_i^{\mathrm{ADEY}}\hat{V}_j^{\mathrm{ADEY}} - V_i^{\mathrm{ADEY}}V_j^{\mathrm{ADEY}}\right)\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\} \\ &\quad + \frac{2}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}(V_i^{\mathrm{ADEY}} - \mathbb{E}[V_i^{\mathrm{ADEY}}])\,\mathbb{E}[V_j^{\mathrm{ADEY}}]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\},\end{aligned} \tag{A.3}$$

where $(\tilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2$ is the infeasible oracle estimator defined as

$$(\tilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2 := \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(V_i^{\mathrm{ADEY}} - \mathbb{E}[V_i^{\mathrm{ADEY}}]\right)\left(V_j^{\mathrm{ADEY}} - \mathbb{E}[V_j^{\mathrm{ADEY}}]\right)\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\}.$$

By Lemma B.4 and Assumptions 4.2, 4.3, and 4.8(i), Proposition 4.1 of Kojevnikov *et al.* (2021) implies that $(\tilde{\sigma}_{S_n}^{\mathrm{ADEY}})^2 = (\sigma_{S_n}^{\mathrm{ADEY}})^2 + o_P(1)$. Thus, we obtain the desired result if the second and third lines of (A.3) are asymptotically negligible.

For the second line of (A.3), observe that

$$\begin{aligned}&\frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(\hat{V}_i^{\mathrm{ADEY}}\hat{V}_j^{\mathrm{ADEY}} - V_i^{\mathrm{ADEY}}V_j^{\mathrm{ADEY}}\right)\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\} \\ &= \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(\hat{V}_i^{\mathrm{ADEY}} - V_i^{\mathrm{ADEY}}\right)\hat{V}_j^{\mathrm{ADEY}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\} \\ &\quad + \frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(\hat{V}_j^{\mathrm{ADEY}} - V_j^{\mathrm{ADEY}}\right)V_i^{\mathrm{ADEY}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\}.\end{aligned} \tag{A.4}$$

Since $\max_{j\in S_n}|\hat{V}_j^{\mathrm{ADEY}}| = O_P(1)$ and $\max_{i\in S_n}|\hat{V}_i^{\mathrm{ADEY}} - V_i^{\mathrm{ADEY}}| = O_P(|S_n|^{-1/2})$ by Assumptions 4.2, 4.3, and 4.8(i) and Lemmas B.1 and B.2, we have

$$\left|\frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\left(\hat{V}_i^{\mathrm{ADEY}} - V_i^{\mathrm{ADEY}}\right)\hat{V}_j^{\mathrm{ADEY}}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) \leqslant b_n\}\right|$$

$$
\begin{aligned}
&\leqslant \left(\max_{j\in S_n}|\widehat{V}_j^{\text{ADEY}}|\right)\left(\max_{i\in S_n}|\widehat{V}_i^{\text{ADEY}}-V_i^{\text{ADEY}}|\right)\frac{1}{|S_n|}\sum_{i\in S_n}\sum_{j\in S_n}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\}\\
&= O_P(1)\cdot O_P\left(\frac{1}{\sqrt{|S_n|}}\right)\cdot M_{S_n}(b_n,1),
\end{aligned}
$$

which is $o_P(1)$ under Assumption 4.8(ii). Similarly, we can show that the second term of (A.4) is $o_P(1)$. Thus, the second line of (A.3) is $o_P(1)$.

To evaluate the third line of (A.3), let $\kappa_{n,i} := \sum_{j\in S_n}\mathbb{E}[V_j^{\text{ADEY}}]\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)\leqslant b_n\}$. Using the norm inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left|\frac{1}{|S_n|}\sum_{i\in S_n}(V_i^{\text{ADEY}}-\mathbb{E}[V_i^{\text{ADEY}}])\kappa_{n,i}\right|\\
&\leqslant \left(\mathbb{E}\left[\left(\frac{1}{|S_n|}\sum_{i\in S_n}(V_i^{\text{ADEY}}-\mathbb{E}\,V_i^{\text{ADEY}})\kappa_{n,i}\right)^2\right]\right)^{1/2}\\
&= \left(\frac{1}{|S_n|^2}\sum_{i\in S_n}\operatorname{Var}[V_i^{\text{ADEY}}]\kappa_{n,i}^2+\frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n\setminus\{i\}}\operatorname{Cov}[V_i^{\text{ADEY}},V_j^{\text{ADEY}}]\kappa_{n,i}\kappa_{n,j}\right)^{1/2}.
\end{aligned}
$$

Noting that V_i^{ADEY} is bounded by Assumptions 4.2 and 4.3, we have $|S_n|^{-2}\sum_{i\in S_n}\operatorname{Var}[V_i^{\text{ADEY}}]\kappa_{n,i}^2 \leqslant C|S_n|^{-1}M_{S_n}(b_n,2)=o(1)$ by Assumption 4.8(ii). Further, we can see that

$$
\begin{aligned}
&\left|\frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n\setminus\{i\}}\operatorname{Cov}[V_i^{\text{ADEY}},V_j^{\text{ADEY}}]\kappa_{n,i}\kappa_{n,j}\right|\\
&\leqslant \frac{1}{|S_n|^2}\sum_{s=1}^{n-1}\sum_{i\in S_n}\sum_{j\in S_n}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)=s\}|\operatorname{Cov}[V_i^{\text{ADEY}},V_j^{\text{ADEY}}]|\cdot|\kappa_{n,i}|\cdot|\kappa_{n,j}|\\
&\leqslant \frac{C}{|S_n|^2}\sum_{s=1}^{n-1}\widetilde{\theta}_{n,s}\sum_{i\in S_n}\sum_{j\in S_n}\sum_{k\in S_n}\sum_{l\in S_n}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)=s\}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,k)\leqslant b_n\}\mathbf{1}\{\ell_{\boldsymbol{A}}(j,l)\leqslant b_n\}\\
&= \frac{C}{|S_n|^2}\sum_{s=1}^{n-1}|\mathcal{J}_{S_n}(s,b_n)|\widetilde{\theta}_{n,s}=o(1),
\end{aligned}
$$

where the second inequality follows from the fact that $\{V_i^{\text{ADEY}}\}_{i\in S_n}$ is ψ-weakly dependent by Lemma B.4 and the last line follows from the second part of Assumption 4.8(i). Thus, the third line of (A.3) is $o_P(1)$. □

B Appendix: Lemmas

B.1 Lemmas for Theorem 4.1

Lemma B.1. Suppose that Assumptions 4.1, 4.4, and 4.6(i) hold. Then, we have

$$\widehat{p}_{S_n}(z,t) - p_{S_n}(z,t) = O_P\left(\frac{1}{\sqrt{|S_n|}}\right)$$

for all $z \in \{0,1\}$ and $t \in \mathcal{T}$.

Proof. By Assumption 4.1(ii), $\mathbb{E}[\widehat{p}_{S_n}(z,t)] = p_{S_n}(z,t)$, and thus it suffices to show that $\mathrm{Var}[\widehat{p}_{S_n}(z,t)] = O(|S_n|^{-1})$. Observe that

$$\begin{aligned}
\mathrm{Var}[\widehat{p}_{S_n}(z,t)] &= \frac{1}{|S_n|^2}\sum_{i\in S_n}\mathrm{Var}[\mathbf{1}\{Z_i=z,T_i=t\}] + \frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n\setminus\{i\}}\mathrm{Cov}[\mathbf{1}\{Z_i=z,T_i=t\},\mathbf{1}\{Z_j=z,T_j=t\}]\\
&= O\left(\frac{1}{|S_n|}\right) + \frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n}\sum_{s\geqslant 1}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)=s\}\,\mathrm{Cov}[\mathbf{1}\{Z_i=z,T_i=t\},\mathbf{1}\{Z_j=z,T_j=t\}]\\
&= O\left(\frac{1}{|S_n|}\right) + \frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n}\sum_{s=1}^{2K}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j)=s\}\,\mathrm{Cov}[\mathbf{1}\{Z_i=z,T_i=t\},\mathbf{1}\{Z_j=z,T_j=t\}],
\end{aligned}$$

where the last equality follows from the fact that, for any $i,j \in S_n$ such that $\ell_{\boldsymbol{A}}(i,j) > 2K$, (Z_i,T_i) is independent of (Z_j,T_j) by Assumptions 4.1(i) and 4.4. By the Cauchy–Schwarz inequality, the second term of the last line is bounded above by $|S_n|^{-1}\sum_{s=1}^{2K} M_{S_n}^{\partial}(s)$ which is $O(|S_n|^{-1})$ by Assumption 4.6(i). This completes the proof. □

Lemma B.2. Suppose that Assumptions 4.1 – 4.6 hold. Then, we have

$$\text{(i)}\quad \widehat{\mu}_{S_n}^{Y}(z,t) - \bar{\mu}_{S_n}^{Y}(z,t) = o_P(1),$$

$$\text{(ii)}\quad \widehat{\mu}_{S_n}^{D}(z,t) - \bar{\mu}_{S_n}^{D}(z,t) = o_P(1),$$

as $|S_n| \to \infty$, for all $z \in \{0,1\}$ and $t \in \mathcal{T}$. Further, $\sqrt{|S_n|}$-consistency is achieved if Assumption 4.6(ii) is strengthened to (4.3).

Proof. We prove only the first result since the second one can be shown in the same way. Observe that

$$\begin{aligned}
\widehat{\mu}_{S_n}^{Y}(z,t) &= \breve{\mu}_{S_n}^{Y}(z,t) - \frac{\breve{\mu}_{S_n}^{Y}(z,t)}{\widehat{p}_{S_n}(z,t)}[\widehat{p}_{S_n}(z,t) - p_{S_n}(z,t)]\\
&= \breve{\mu}_{S_n}^{Y}(z,t) + O_P\left(\frac{1}{\sqrt{|S_n|}}\right),
\end{aligned}$$

by Lemma B.1 and Assumptions 4.2 and 4.3. Here, it is easy to see that $\mathbb{E}[\breve{\mu}_{S_n}^{Y}(z,t)] = \bar{\mu}_{S_n}^{Y}(z,t)$. Further, letting $Q_i^Y := Y_i \mathbf{1}\{Z_i = z, T_i = t\}/p_{S_n}(z,t)$, we can see that

$$\begin{aligned}\mathrm{Var}\left[\breve{\mu}^Y_{S_n}(z,t)\right] &= \frac{1}{|S_n|^2}\sum_{i\in S_n}\mathrm{Var}[Q_i^Y] + \frac{1}{|S_n|^2}\sum_{i\in S_n}\sum_{j\in S_n\backslash\{i\}}\mathrm{Cov}[Q_i^Y, Q_j^Y] \\ &= O\left(\frac{1}{|S_n|}\right) + \frac{1}{|S_n|^2}\sum_{s=1}^{n-1}\sum_{i\in S_n}\sum_{j\in S_n}\mathbf{1}\{\ell_{\boldsymbol{A}}(i,j) = s\}\,\mathrm{Cov}[Q_i^Y, Q_j^Y].\end{aligned}$$

Using Assumptions 4.1–4.5, similar arguments to the proof of Theorem 2 of Leung (2021) can show that the second term in the last line is bounded above by $C|S_n|^{-1}\sum_{s=1}^{n-1} M^{\partial}_{S_n}(s)\widetilde{\theta}_{n,s}$ for some positive constant C. Thus, we obtain the desired result by Assumption 4.6(ii) or (4.3) and Chebyshev's inequality. □

B.2 Lemmas for Theorem 4.2

For completeness, we define ψ-dependence in line with Definition 2.2 of Kojevnikov *et al.* (2021). For $d \in \mathbb{N}$, let $\mathcal{L}_d$ be the set of real-valued bounded Lipschitz functions on $\mathbb{R}^d$:

$$\mathcal{L}_d := \{f : \mathbb{R}^d \to \mathbb{R} : \ \|f\|_\infty < \infty,\ \mathrm{Lip}(f) < \infty\},$$

where $\|f\|_\infty := \sup_{x\in\mathbb{R}^d}|f(x)|$ and $\mathrm{Lip}(f)$ indicates the Lipschitz constant of f (with respect to the Euclidean norm). We write the distance between subsets $H, H' \subset S_n$ by $\ell_{\boldsymbol{A}}(H,H') := \min\{\ell_{\boldsymbol{A}}(i,j) : i \in H, j \in H'\}$. For $h, h' \in \mathbb{N}$, denote the collection of pairs (H, H') whose sizes are h and h', respectively, with distance at least s as

$$\mathcal{P}_{S_n}(h,h',s) := \{(H,H') : H, H' \subset S_n, |H| = h, |H'| = h', \ell_{\boldsymbol{A}}(H,H') \geqslant s\}.$$

For a generic random vector $\boldsymbol{W}_{n,i} \in \mathbb{R}^v$, let $\boldsymbol{W}_{n,H} = (\boldsymbol{W}_{n,i})_{i\in H}$ and $\boldsymbol{W}_{n,H'} = (\boldsymbol{W}_{n,i})_{i\in H'}$.

Definition B.1 (ψ-dependence)**.** A triangular array $\{\boldsymbol{W}_{n,i}\}_{i\in S_n}$ is called *ψ-dependent*, if for each $n \in \mathbb{N}$, there exist a sequence of uniformly bounded constants $\{\widetilde{\theta}_{n,s}\}_{s\geqslant 0}$ with $\widetilde{\theta}_{n,0} = 1$ and a collection of nonrandom functions $\{\psi_{h,h'}\}_{h,h'\in\mathbb{N}}$, where $\psi_{h,h'} : \mathcal{L}_{hv} \times \mathcal{L}_{h'v} \to [0,\infty)$, such that for all $s > 0$, $(H,H') \in \mathcal{P}_{S_n}(h,h',s)$, $f \in \mathcal{L}_{hv}$, and $f' \in \mathcal{L}_{h'v}$,

$$|\,\mathrm{Cov}[f(\boldsymbol{W}_{n,H}), f'(\boldsymbol{W}_{n,H'})]| \leqslant \psi_{h,h'}(f,f')\widetilde{\theta}_{n,s}.$$

The sequence $\{\widetilde{\theta}_{n,s}\}_{s\geqslant 0}$ is called the *dependence coefficients* of $\{\boldsymbol{W}_{n,i}\}_{i\in S_n}$. Further, if $\sup_{n\in\mathbb{N}}\widetilde{\theta}_{n,s} \to 0$ as $s \to \infty$, we say that $\{\boldsymbol{W}_{n,i}\}_{i\in S_n}$ is *ψ-weakly dependent*.

Denote $\boldsymbol{W}_i := (W_i^Y, W_i^D, W_i^Z, W_i^{1-Z})$, whose elements are as defined in (4.4). For a subset $H \subset S_n$ with $|H| = h$, we write $\boldsymbol{W}_H = (\boldsymbol{W}_i)_{i\in H}$.

Lemma B.3. Under Assumptions 4.1 – 4.5, the triangular array $\{\boldsymbol{W}_i\}_{i\in S_n}$ is ψ-weakly dependent with the dependence coefficients $\{\widetilde{\theta}_{n,s}\}_{s\geqslant 0}$ defined by (4.2) and

$$\psi_{h,h'}(f,f') = C[\|f\|_\infty\|f'\|_\infty + h\|f'\|_\infty\,\mathrm{Lip}(f) + h'\|f\|_\infty\,\mathrm{Lip}(f')], \qquad \forall h,h' \in \mathbb{N}, f \in \mathcal{L}_{4h}, f' \in \mathcal{L}_{4h'},$$

with some positive constant C.

Proof. Consider arbitrary $n, h, h' \in \mathbb{N}$, $s > 0$, $(H, H') \in \mathcal{P}_{S_n}(h, h'; s)$, $f \in \mathcal{L}_{4h}$, and $f' \in \mathcal{L}_{4h'}$. Let $\xi := f(\boldsymbol{W}_H)$ and $\zeta := f'(\boldsymbol{W}_{H'})$. Consider two independent copies of $\boldsymbol{Z}$, say $\boldsymbol{Z}'$ and $\boldsymbol{Z}''$. For $i \in H$ and $j \in H'$, define $\boldsymbol{Z}_i^{(s,\xi)} := (\boldsymbol{Z}_{N_{\boldsymbol{A}}(i,s)}, \boldsymbol{Z}'_{N^c_{\boldsymbol{A}}(i,s)})$, $\boldsymbol{Z}_j^{(s,\zeta)} := (\boldsymbol{Z}_{N_{\boldsymbol{A}}(j,s)}, \boldsymbol{Z}''_{N^c_{\boldsymbol{A}}(j,s)})$, and

$$W_i^{Y,(s,\xi)} := y_i(\boldsymbol{Z}_i^{(s,\xi)}) \left[\frac{\mathbf{1}\{Z_i = 1, T(i, \boldsymbol{Z}_i^{(s,\xi)}, \boldsymbol{A}) = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i = 0, T(i, \boldsymbol{Z}_i^{(s,\xi)}, \boldsymbol{A}) = t\}}{p_{S_n}(0,t)} \right],$$
$$W_j^{Y,(s,\zeta)} := y_j(\boldsymbol{Z}_j^{(s,\zeta)}) \left[\frac{\mathbf{1}\{Z_j = 1, T(j, \boldsymbol{Z}_j^{(s,\zeta)}, \boldsymbol{A}) = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_j = 0, T(j, \boldsymbol{Z}_j^{(s,\zeta)}, \boldsymbol{A}) = t\}}{p_{S_n}(0,t)} \right].$$

We similarly define $W_i^{D,(s,\xi)}$, $W_j^{D,(s,\zeta)}$, $W_i^{Z,(s,\xi)}$, $W_j^{Z,(s,\zeta)}$, $W_i^{1-Z,(s,\xi)}$, and $W_j^{1-Z,(s,\zeta)}$, and let

$$\boldsymbol{W}_i^{(s,\xi)} := (W_i^{Y,(s,\xi)}, W_i^{D,(s,\xi)}, W_i^{Z,(s,\xi)}, W_i^{1-Z,(s,\xi)}), \qquad \boldsymbol{W}_H^{(s,\xi)} := (\boldsymbol{W}_i^{(s,\xi)})_{i\in H}, \qquad \xi^{(s)} := f(\boldsymbol{W}_H^{(s,\xi)}),$$
$$\boldsymbol{W}_j^{(s,\zeta)} := (W_j^{Y,(s,\zeta)}, W_j^{D,(s,\zeta)}, W_j^{Z,(s,\zeta)}, W_j^{1-Z,(s,\zeta)}), \qquad \boldsymbol{W}_{H'}^{(s,\zeta)} := (\boldsymbol{W}_j^{(s,\zeta)})_{j\in H'}, \qquad \zeta^{(s)} := f'(\boldsymbol{W}_{H'}^{(s,\zeta)}).$$

Since f and f' are bounded functions,

$$\begin{aligned} |\operatorname{Cov}(\xi,\zeta)| &= |\operatorname{Cov}(\xi,\zeta)| \cdot \mathbf{1}\{s \leqslant 2\max\{K,1\}\} + |\operatorname{Cov}(\xi,\zeta)| \cdot \mathbf{1}\{s > 2\max\{K,1\}\} \\ &\leqslant 2\|f\|_\infty \|f'\|_\infty \cdot \mathbf{1}\{s \leqslant 2\max\{K,1\}\} + |\operatorname{Cov}(\xi,\zeta)| \cdot \mathbf{1}\{s > 2\max\{K,1\}\}. \end{aligned}$$

For the second term, recall that $\ell_{\boldsymbol{A}}(H, H') > 2\max\{K,1\}$ when $s > 2\max\{K,1\}$. Then, denoting $s' = \lfloor s/2 \rfloor$, Assumptions 4.1(i) and 4.4 imply that $\boldsymbol{W}_H^{(s',\xi)}$ is independent of $\boldsymbol{W}_{H'}^{(s',\zeta)}$. From this, we have

$$\begin{aligned} |\operatorname{Cov}(\xi,\zeta)| &\leqslant |\operatorname{Cov}(\xi - \xi^{(s')}, \zeta)| + |\operatorname{Cov}(\xi^{(s')}, \zeta - \zeta^{(s')})| + |\operatorname{Cov}(\xi^{(s')}, \zeta^{(s')})| \\ &= |\operatorname{Cov}(\xi - \xi^{(s')}, \zeta)| + |\operatorname{Cov}(\xi^{(s')}, \zeta - \zeta^{(s')})| \\ &\leqslant 2\|f'\|_\infty \mathbb{E}\,|\xi - \xi^{(s')}| + 2\|f\|_\infty \mathbb{E}\,|\zeta - \zeta^{(s')}| \\ &\leqslant 2\|f'\|_\infty \operatorname{Lip}(f)\, \mathbb{E}\,\|\boldsymbol{W}_H - \boldsymbol{W}_H^{(s',\xi)}\| + 2\|f\|_\infty \operatorname{Lip}(f')\, \mathbb{E}\,\|\boldsymbol{W}_{H'} - \boldsymbol{W}_{H'}^{(s',\zeta)}\|, \end{aligned}$$

where $\|\cdot\|$ denotes the Euclidean norm. Here, by Assumption 4.4,

$$W_i^Y - W_i^{Y,(s',\xi)} = [y_i(\boldsymbol{Z}) - y_i(\boldsymbol{Z}_i^{(s',\xi)})] \left(\frac{\mathbf{1}\{Z_i = 1, T(i, \boldsymbol{Z}, \boldsymbol{A}) = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i = 0, T(i, \boldsymbol{Z}, \boldsymbol{A}) = t\}}{p_{S_n}(0,t)} \right)$$

and

$$W_i^D - W_i^{D,(s',\xi)} = [D_i(\boldsymbol{Z}) - D_i(\boldsymbol{Z}_i^{(s',\xi)})] \left(\frac{\mathbf{1}\{Z_i = 1, T(i, \boldsymbol{Z}, \boldsymbol{A}) = t\}}{p_{S_n}(1,t)} - \frac{\mathbf{1}\{Z_i = 0, T(i, \boldsymbol{Z}, \boldsymbol{A}) = t\}}{p_{S_n}(0,t)} \right).$$

Further, it is easy to see that $W_i^Z - W_i^{Z,(s',\xi)} = 0$ and $W_i^{1-Z} - W_i^{1-Z,(s',\xi)} = 0$ by Assumption 4.4. Thus, by Assumptions 4.2 and 4.3, $\mathbb{E}\,\|\boldsymbol{W}_H - \boldsymbol{W}_H^{(s',\xi)}\| \leqslant Ch\theta_{n,s'}$ for some positive constant C. In the same way, we can see that $\mathbb{E}\,\|\boldsymbol{W}_{H'} - \boldsymbol{W}_{H'}^{(s',\zeta)}\| \leqslant Ch'\theta_{n,s'}$. In conjunction with Assumption 4.5, this completes the proof. □

The next lemma is immediate from Lemma B.3 (cf. Lemma 2.1 of Kojevnikov *et al.*, 2021). Let $\{\boldsymbol{c}_{n,i}\}_{i \in S_n}$ be a sequence of uniformly bounded nonrandom vectors in $\mathbb{R}^4$.

Lemma B.4. Under Assumptions 4.1 – 4.5, the triangular array $\{\boldsymbol{c}_{n,i}^\top \boldsymbol{W}_i\}_{i \in S_n}$ is ψ-weakly dependent with the dependence coefficients $\{\widetilde{\theta}_{n,s}\}_{s \geqslant 0}$ defined by (4.2) and

$$\psi_{h,h'}(f, f') = C[\|f\|_\infty \|f'\|_\infty + h\|f'\|_\infty \operatorname{Lip}(f) + h'\|f\|_\infty \operatorname{Lip}(f')], \qquad \forall h, h' \in \mathbb{N}, f \in \mathcal{L}_h, f' \in \mathcal{L}_{h'},$$

with some positive constant C.

References

Abadie, A., Athey, S., Imbens, G.W., and Wooldridge, J.M., 2020. Sampling-based versus design-based uncertainty in regression analysis, *Econometrica*, 88 (1), 265–296.

Angrist, J.D., Imbens, G.W., and Rubin, D.B., 1996. Identification of causal effects using instrumental variables, *Journal of the American Statistical Association*, 91 (434), 444–455.

Aronow, P.M., Eckles, D., Samii, C., and Zonszein, S., 2021. Spillover effects in experimental data, *Advances in Experimental Political Science*, 289–319.

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 (4), 1912–1947.

Athey, S., Eckles, D., and Imbens, G.W., 2018. Exact p-values for network interference, *Journal of the American Statistical Association*, 113 (521), 230–240.

Basse, G., Feller, A., and Toulis, P., 2019. Randomization tests of causal effects under interference, *Biometrika*, 106 (2), 487–494.

Bramoullé, Y., Djebbari, H., and Fortin, B., 2009. Identification of peer effects through social networks, *Journal of Econometrics*, 150 (1), 41–55.

Chin, A., 2018. Central limit theorems via stein's method for randomized experiments under interference, *arXiv preprint arXiv:1804.03105*.

DiTraglia, F.J., Garcia-Jimeno, C., O'Keeffe-O'Donovan, R., and Sánchez-Becerra, A., 2021. Identifying causal effects in experiments with spillovers and non-compliance, *arXiv preprint arXiv:2011.07051*.

Dupas, P., 2014. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment, *Econometrica*, 82 (1), 197–228.

Egami, N., 2021. Spillover effects in the presence of unobserved networks, *Political Analysis*, 29 (3), 287–316.

Fisher, R.A., 1935. *The design of experiments*, London: Oliver & Boyd.

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 (534), 901–918.

Halloran, M.E. and Hudgens, M.G., 2016. Dependent happenings: a recent methodological review, *Current Epidemiology Reports*, 3 (4), 297–305.

Heckman, J.J. and Pinto, R., 2018. Unordered monotonicity, *Econometrica*, 86 (1), 1–35.

Hong, G. and Raudenbush, S.W., 2006. Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data, *Journal of the American Statistical Association*, 101 (475), 901–910.

Hudgens, M.G. and Halloran, M.E., 2008. Toward causal inference with interference, *Journal of the American Statistical Association*, 103 (482), 832–842.

Imai, K., Jiang, Z., and Malani, A., 2021. Causal inference with interference and noncompliance in two-stage randomized experiments, *Journal of the American Statistical Association*, 116 (534), 632–644.

Imbens, G.W. and Angrist, J.D., 1994. Identification and estimation of local average treatment effects, *Econometrica*, 62 (2), 467–475.

Imbens, G.W. and Rubin, D.B., 2015. *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press.

Kang, H. and Imbens, G., 2016. Peer encouragement designs in causal inference with partial interference and identification of local average network effects, *arXiv preprint arXiv:1609.04464*.

Kojevnikov, D., 2021. The bootstrap for network dependent processes, *arXiv preprint arXiv:2101.12312*.

Kojevnikov, D., Marmer, V., and Song, K., 2021. Limit theorems for network dependent random variables, *Journal of Econometrics*, 222 (2), 882–908.

Leung, M.P., 2021. Causal inference under approximate neighborhood interference, *Econometrica*, forthcoming.

Li, X., Ding, P., Lin, Q., Yang, D., and Liu, J.S., 2019. Randomization inference for peer effects, *Journal of the American Statistical Association*, 114 (528), 1651–1664.

Manski, C.F., 2013. Identification of treatment response with social interactions, *The Econometrics Journal*, 16 (1), S1–S23.

Miguel, E. and Kremer, M., 2004. Worms: identifying impacts on education and health in the presence of treatment externalities, *Econometrica*, 72 (1), 159–217.

Paluck, E.L., Shepherd, H., and Aronow, P.M., 2016. Changing climates of conflict: A social network experiment in 56 schools, *Proceedings of the National Academy of Sciences*, 113 (3), 566–571.

Paluck, E.L., Shepherd, H.R., and Aronow, P., 2020. Changing climates of conflict: A social network experiment in 56 schools, New Jersey, 2012-2013, Inter-university Consortium for Political and Social Research [distributor], 2020-09-14. https://doi.org/10.3886/ICPSR37070.v2.

Rubin, D.B., 1980. Discussion of "randomization analysis of experimental data in the fisher randomization test" by D. Basu, *Journal of the American Statistical Association*, 75, 591–593.

Sävje, F., 2021. Causal inference with misspecified exposure mappings, *arXiv preprint arXiv:2103.06471*.

Sävje, F., Aronow, P.M., and Hudgens, M.G., 2021. Average treatment effects in the presence of unknown interference, *The Annals of Statistics*, 49 (2), 673–701.

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 (476), 1398–1407.

VanderWeele, T.J. and An, W., 2013. Social networks and causal inference, *Handbook of Causal Analysis for Social Research*, 353–374.

Vazquez-Bare, G., 2021. Causal spillover effects using instrumental variables, *arXiv preprint arXiv:2003.06023*.

Zelizer, A., 2019. Is position-taking contagious? evidence of cue-taking from two field experiments in a state legislature, *American Political Science Review*, 113 (2), 340–352.