Original Articles

Sociological Methodology
1–16

DOI: 10.1177/0081175018782569
http://sm.sagepub.com

LIMITATIONS OF DESIGN-BASED CAUSAL INFERENCE AND A/B TESTING UNDER ARBITRARY AND NETWORK INTERFERENCE

Guillaume W. Basse*
Edoardo M. Airoldi†

Abstract

Randomized experiments on a network often involve interference between connected units, namely, a situation in which an individual's treatment can affect the response of another individual. Current approaches to deal with interference, in theory and in practice, often make restrictive assumptions on its structure—for instance, assuming that interference is local—even when using otherwise nonparametric inference strategies. This reliance on explicit restrictions on the interference mechanism suggests a shared intuition that inference is impossible without any assumptions on the interference structure. In this paper, we begin by formalizing this intuition in the context of a classical nonparametric approach to inference, referred to as design-based inference of causal effects. Next, we show how, always in the context of design-based inference, even parametric structural assumptions that allow the existence of unbiased estimators cannot guarantee a decreasing variance even in the large sample limit. This lack of concentration in large samples is often observed empirically, in randomized experiments in which interference of some form is expected to

*University of California-Berkeley, CA
†Temple University, Philadelphia, PA

Corresponding Author:
Edoardo M. Airoldi, Department of Statistical Science, Fox School of Business, Temple University, 1810 Liacouras Walk, Philadelphia, PA 19122, USA
Email: airoldi@alumni.harvard.edu

be present. This result has direct consequences for the design and analysis of large experiments—for instance, in online social platforms—where the belief is that large sample sizes automatically guarantee small variance. More broadly, our results suggest that although strategies for causal inference in the presence of interference borrow their formalism and main concepts from the traditional causal inference literature, much of the intuition from the no-interference case do not easily transfer to the interference setting.

Keywords

causal inference, social network data, interference

1. INTRODUCTION

In modern randomized experiments, *interference* typically means that the response of a given unit to a certain treatment may depend on the treatment assigned to other units. In online marketing, Aral and Walker (2012) showed that adoption of a product by an individual in a social network tends to encourage the adoption of the same product by their neighbors. Similar examples can be found in political science where encouraging an individual to vote is expected to increase the turnout for other members of the household (Sinclair, McConnell, and Green 2012). Researchers have had some success in addressing the question by making assumptions on the interference mechanism (Basse and Feller forthcoming; Bowers et al. 2018; Bowers, Fredrickson, and Aronow 2016; Ugander et al. 2013). However, because these mechanisms are often very complex, it is tempting to believe that we can rely on large sample sizes to avoid making assumptions. The goal of this paper is to show how wrong this intuition is, precisely.

1.1. *Background and Set-up*

In a randomized experiment (or A/B test), N units are randomly assigned to treatments A or B, and an outcome of interest $Y_i(\mathbf{Z}^{obs})$ is measured for each unit i, where $\mathbf{Z}^{obs} \in \mathcal{Z} = \{A, B\}^N$ is the observed assignment vector. There are generally two paradigms available for making inference: the model-based approach and the design-based approach.

In the model-based approach, the vector of observed outcomes is modeled $Y_i(\mathbf{Z}^{obs}) \sim F(\vec{\xi})$. The estimands are generally functions of the parameters $\theta = g(\vec{\xi})$, and the estimators are usually obtained using maximum likelihood or Bayesian inference. Two key points are that the estimators are selected using the model $F(\vec{\xi})$ and the properties of the estimator are typically derived in some asymptotic regime and incorporate randomness due to both the model for the outcomes and the randomization distribution.

In contrast, the design-based approach we consider in this paper takes the opposite perspective and considers the *potential outcomes* $\{\mathbf{Y}(\mathbf{Z})\}_{\mathbf{Z}}$ as fixed but a priori unknown quantities. The assignment mechanism $\mathcal{R}$, which assigns a probability $\mathbb{P}_{\mathcal{R}}(\mathbf{Z})$ to every vector $\mathbf{Z} \in \mathcal{Z}$, provides the only source of randomness. Estimands are then generally functions of the potential outcomes $\theta = g(\{\mathbf{Y}(\mathbf{Z})\})$, and it is desirable to find a pair $(\hat{\theta}, \mathbb{P}_{\mathcal{R}})$ such that the estimator $\hat{\theta}$ has good properties under the design $\mathcal{R}$ in finite samples. One goal of this paper is to show that under arbitrary interference, there exist no pair $(\hat{\theta}, \mathbb{P}_{\mathcal{R}})$ with good properties.

A third approach to inference exists, typically referred to as *model-assisted*, which attempts to blend the two approaches by using a model for the outcomes to inform the choice of design or correction factors for estimators in a given family (Basse and Airoldi forthcoming; Särndal, Swensson, and Wretman 2003). Since the actual inference is then performed from a design-based perspective, this approach also falls under the scope of our paper.

The idea that inference is not feasible without assumptions on the interference mechanism has been suggested in the literature for observational studies (e.g., see Shalizi and Thomas 2011) and also in the context of randomized experiments; for instance, Aronow and Samii (forthcoming) state that under arbitrary interference, "it is clear that there would be no meaningful way to use the results of the experiment" and then focus on how to do inference under some form of restricted interference. In the first, and what is to our knowledge also the only, attempt at formalizing the issue, Manski (2013) shows that in a model-based setting, the distribution of potential outcomes is not identifiable under arbitrary interference.

1.2. *Contributions*

Our paper makes three contributions. First, we extend and clarify the results of Manski (2013) in the context of design-based inference (Section 2) and prove, among other things, that there exist no consistent estimators under arbitrary interference. Second, we extend the work of

Aronow and Samii (forthcoming) by focusing on a popular design and an interference structure commonly assumed in network settings. Assuming an Erdos-Renyi for the network, we show that for a class of unbiased estimators, consistency depends on the parameter p of the model for the graph. Finally, using the concept of *effective treatment* introduced by Manski (2013), we provide analytical insights into the general problem of interference and the convergence of estimators.

For methodologists, this paper is meant to provide a clear formalization of some intuitions that have been suggested. For practitioners, this paper is meant to offer a convincing argument for the necessity of making explicit assumptions about the structure of the interference mechanism to be able to rely on asymptotic standard errors for causal inference.

2. NO INTERFERENCE AND ARBITRARY INTERFERENCE

2.1. *Set-up*

To avoid pathological cases, we assume that the potential outcomes are bounded:

Assumption 1 (bounded outcomes): There exists M such that for all N:

$$0 < Y_i(\boldsymbol{Z}) < M \quad \forall\, i = 1, \ldots N$$

Although many of our results could generalize to large classes of designs $\mathcal{R}$, we will focus on three assignment mechanisms for the sake of clarity. The completely randomized design (CRD) assigns a fixed number N_A of randomly selected units to treatment A and the other $N_B = N - N_A$ units to treatment B. The Bernoulli design (BD) assigns independently each unit to treatments A or B with probability half, while the conditional Bernoulli design (CBD) operates similarly to the BD but excludes the assignments $\mathbf{Z} = \mathbf{A}$ and $\mathbf{Z} = \mathbf{B}$ in which all units are assigned to A or all units assigned to B.

The estimands considered in this paper are of the form $\theta(\mathbf{A}, \mathbf{B}) = g(\mathbf{Y}(\mathbf{A}), \mathbf{Y}(\mathbf{B}))$, and estimators will be denoted by $\hat{\theta}(\mathbf{Z})$.

Example 1. The most popular estimand in causal inference is the average total treatment effect (Imbens and Rubin 2015):

$$\theta(\mathbf{A}, \mathbf{B}) = \overline{Y}(\mathbf{A}) - \overline{Y}(\mathbf{B}),$$

but our results are stated in greater generality.

All statements about estimators must hold regardless of the value of the potential outcomes $\{Y(\mathbf{Z})\}_{\mathbf{Z}}$. In particular, we consider the following definition of unbiasedness.

Definition 1. Let $\theta(\mathbf{A}, \mathbf{B}) = g(\mathbf{Y}(\mathbf{A}), \mathbf{Y}(\mathbf{B}))$ be the estimand of interest. An estimator $\hat{\theta}(Z)$ is said to be unbiased if

$$\mathbb{E}_{\mathcal{R}}[\hat{\theta}(\mathbf{Z})] = \theta(\mathbf{A}, \mathbf{B})$$

for all values of the potential outcomes $\mathbf{Y}(\mathbf{Z})\}_{\mathbf{Z}}$ (the expectation being taken with respect to the assignment mechanisms).

In particular, an estimator is not unbiased if the equality $E(\hat{\theta}(\mathbf{Z})) = \theta(\mathbf{A}, \mathbf{B})$ only holds for specific values of $\{\mathbf{Y}(\mathbf{Z})\}_{Z}$. Similar considerations apply to the concept of consistency.

2.2. *Inference under No Interference*

The most prominent applications involve situations in which interference is a nuisance and can be safely assumed away. The main clause of the popular stable unit treatment value assumption (SUTVA; Rubin 1980) implies that there is no interference between units:

Assumption 2 (no interference):

$$Y_i(\mathbf{Z}) = Y_i(Z_i) \quad \forall\, i = 1, \ldots, N$$

Studying the no interference case is important because it often implicitly guides our intuition in more complex scenarios, especially the large sample behavior or estimators. One aim of this paper is to show the dangers of that intuition.

We illustrate the standard characteristics of inference under this assumption by focusing on the CRD and the average total treatment effect:

$$\theta(\mathbf{A}, \mathbf{B}) = \frac{1}{N} \sum_{i}^{N} (Y_i(\mathbf{A}) - Y_i(\mathbf{B}))$$

as our estimand. The following estimator:

$$\hat{\theta}(\mathbf{Z}) = \frac{1}{N_A}\sum_{i}^{N} I(Z_i = A)Y_i(\mathbf{Z}) - \frac{1}{N_B}\sum_{i}^{N} I(Z_i = B)Y_i(\mathbf{Z}), \tag{1}$$

which under Assumption 2 simplifies to

$$\hat{\theta}(\mathbf{Z}) = \frac{1}{N_A}\sum_{i}^{N} I(Z_i = A)Y_i(A_i) - \frac{1}{N_B}\sum_{i}^{N} I(Z_i = B)Y_i(B_i),$$

can easily be shown to be unbiased for $\theta(\mathbf{A}, \mathbf{B})$ under the CRD. That is

$$\mathbb{E}_{\mathcal{R}}[\hat{\theta}(\mathbf{Z})] = \theta(\mathbf{A}, \mathbf{B}),$$

and the variance is (see e.g., Imbens and Rubin 2015, chapter 6)

$$\begin{aligned}\mathbb{V}_{\mathcal{R}}[\hat{\theta}] = \tfrac{V_A}{N_A} + \tfrac{V_B}{N_B} - \tfrac{V_\theta}{N} &\leq \tfrac{1}{N_A}\tfrac{1}{N-1}\sum_i (Y_i(\mathbf{A}) - \overline{Y}(\mathbf{A}))^2 + \tfrac{1}{N_B}\tfrac{1}{N-1}\sum_i (Y_i(\mathbf{B}) - \overline{Y}(\mathbf{B}))^2 \\ &\leq \tfrac{1}{N_A(N-1)}NM^2 + \tfrac{1}{N_B(N-1)}NM^2 \\ &\leq \tfrac{4M^2}{N-1},\end{aligned}$$

where the terms V_A, V_B, and V_θ are defined in the supplemental online appendix. So under Assumption 1 and Assumption 2, there exists an unbiased estimator, with variance of order $O(1/N)$.

2.3. *Arbitrary Interference*

Under arbitrary interference, the outcome for unit i depends on the entire assignment vector $\mathbf{Z}$, not just on its own assignment Z_i. We show in this section that not only is the estimator in Equation 1 biased but that unbiased estimators (in the sense of Definition 1) simply do not exist for a wide class of designs, which includes both the CRD and the CBD.

Theorem 1. Consider any nondegenerate[1] estimand $\theta(\mathbf{A}, \mathbf{B})$ and any assignment mechanism $\mathcal{R}$ such that $\mathbb{P}_{\mathcal{R}}(\mathbf{Z} = \mathbf{A}) = \mathbb{P}_{\mathcal{R}}(\mathbf{Z} = \mathbf{B}) = 0$. There exists no unbiased estimator of $\theta(\mathbf{A}, \mathbf{B})$ under $\mathcal{R}$.

If the design assigns non-zero probability to the treatment allocation vectors $\mathbf{A}$ and $\mathbf{B}$, then unbiased estimators exist for a restricted class of estimands. We formalize this idea in the case of the BD in Proposition

1 but show in Theorem 2 that no estimator can have both low bias and low variance for all values of the potential outcomes.

Proposition 1. Consider the BD. If the estimand is of additive form $\theta(\mathbf{A},\mathbf{B}) = \theta_1(\mathbf{A}) + \theta_2(\mathbf{B})$, then unbiased estimators are of the form

$$\hat{\theta}(\mathbf{Z}) = C(\mathbf{Z}) + 2^N I(\mathbf{Z}=\mathbf{A})\theta_1(\mathbf{A}) + 2^N I(\mathbf{Z}=\mathbf{B})\theta_2(\mathbf{B}),$$

where $C(\mathbf{Z})$ does not depend on any potential outcomes[2] and satisfies

$$\sum_{Z\in\mathcal{Z}} C(\mathbf{Z}) = 0.$$

For other types of estimands $\theta(\mathbf{A},\mathbf{B})$, there exist no unbiased estimators.

Taken together, Theorem 1 and Proposition 1 formalize a simple idea: If $\mathbf{Z} \neq \mathbf{Z}'$, then $\mathbf{Y}(\mathbf{Z}')$ is completely noninformative for $\mathbf{Y}(\mathbf{Z})$. In particular, the only assignments that, if observed, could provide information about the estimand $\theta(\mathbf{A},\mathbf{B})$ are $\mathbf{Z}=\mathbf{A}$ and $\mathbf{Z}=\mathbf{B}$. The estimator of Proposition 1 reflects this by evaluating to 0 for any uninformative assignment (i.e., $\mathbf{Z} \notin \{\mathbf{A},\mathbf{B}\}$) and assigning large weight to the only two informative assignments ($\mathbf{Z}=\mathbf{A}$ and $\mathbf{Z}=\mathbf{B}$). This is a known case of failure for this kind of Horvitz-Thompson estimators (Basu 2011). The following example shows the implications of Theorem 1 and Proposition 1 when the estimand is the average total treatment effect.

Example 2. Consider the following estimand:

$$\theta(\mathbf{A},\mathbf{B}) = \overline{Y}(\mathbf{A}) - \overline{Y}(\mathbf{B})$$

and let $0 < N_1 < N$. $\theta(\mathbf{A},\mathbf{B})$ has the additive form of Proposition 1 with $\theta_1(\mathbf{A}) = \overline{Y}(\mathbf{A})$ and $\theta_2(\mathbf{B}) = -\overline{Y}(\mathbf{B})$, so under BD and in the absence of external information (i.e., setting $C(\mathbf{Z}) = 0$ for all $\mathbf{Z}$), the only unbiased estimator of $\theta(\mathbf{A},\mathbf{B})$ is

$$\hat{\theta}(Z) = \begin{cases} 2^N\overline{Y}(\mathbf{A}) & \textit{if } \mathbf{Z}=\mathbf{A} \\ -2^N\overline{Y}(\mathbf{B}) & \textit{if } \mathbf{Z}=\mathbf{B} \\ 0 & \textit{otherwise}. \end{cases}$$

Theorem 1, however, states there exist no unbiased estimators for this estimand under CRD and CBD, both of which assign zero probability to assignments **A** and **B**.

Finite sample bias is not uncommon in statistical applications and is generally acceptable if it can be traded for a large reduction in variance and vanishes as the sample size increase. The next theorem studies this tradeoff for both the BD and CRD by looking at the mean square error (MSE).

Theorem 2. Consider any estimand $\theta(\mathbf{A},\mathbf{B}) = g(\mathbf{Y}(\mathbf{A}),\mathbf{Y}(\mathbf{B}))$, and suppose that g is onto[3] [0, M]. Then for all sample size N and estimator $\hat{\theta}(\mathbf{Z})$, there exist potential outcomes satisfying Assumption 1 such that

$$MSE(\hat{\theta},\theta) \geq \frac{M^2}{8},$$

where the MSE is taken under CRD or BD.

In the previous section, we showed that under Assumption 2 there exists an estimator for the average total treatment effect that was unbiased, had a variance of order $O(1/N)$, and thus an MSE of order $O(1/N)$. These properties of the estimator hold for all possible values of the potential outcomes. Theorem 2 states that under arbitrary interference, this is no longer the case: Whether one is conducting a small study or a large-scale experiment on a social network, no estimator will perform well for all values of the potential outcomes. A direct consequence of the theorem is that there exists no consistent estimator under arbitrary interference.

Remark 1. The condition on g in Theorem 2 can be relaxed without affecting the main idea behind the theorem. It is useful to note that the difference in means estimand $\theta(\mathbf{A},\mathbf{B}) = \overline{Y}(\mathbf{A}) - \overline{Y}(\mathbf{B})$ satisfies this condition.

3. CAUTION WHEN STRUCTURING THE INTERFERENCE MECHANISM

Assumptions restricting the interference mechanism (usually involving a notion of locality) can alleviate some of the issues mentioned in the previous section (Basse and Feller forthcoming; Hudgens and Halloran

2008; Ugander et al. 2013) and allow the existence of unbiased estimators for a variety of estimands (Aronow and Samii forthcoming; Ugander et al. 2013). Even then, however, basic intuition from the no interference case remains misleading. When considering such assumptions, it is important to assess their impact on both the bias and the variance of estimators, especially in settings that naturally afford large sample sizes, where it is tempting to focus mostly on the bias, and ignore potential issues with the variance. We illustrate this danger with a realistic example in which a local interference assumption leads to an estimator that is unbiased but whose variance explodes in large sample settings (Aronow and Samii forthcoming; Ugander et al. 2013).

If $d(i,j)$ is a measure of the distance between units i and j, it is often plausible to assume that only units in the *k-step neighborhood* $\mathcal{N}_i^{(k)} = \{j = 1 \ldots N : d(i,j) \leq k\}$ of i can interfere with its outcome (Coppock and Sircar 2013; Ugander et al. 2013). Formally,

Assumption 3 (k-local interference):

$$Y_i(\mathbf{Z}) = Y_i(\mathbf{Z}_{\mathcal{N}_i^{(k)}}),$$

where $\mathbf{Z}_{\mathcal{N}_i^{(k)}}$ denotes the sub vector of $\mathbf{Z}$ containing the assignments of the k-step neighbors of unit i. This assumption gives a special role to the following subsets of $\mathcal{Z}$:

$$\mathcal{Z}_i^{(k)}(\mathbf{A}) = \{\mathbf{Z} : \mathbf{Z}_{\mathcal{N}_i^{(k)}} = \mathbf{A}_{\mathcal{N}_i^{(k)}}\}, \tag{2}$$

where $\mathcal{Z}_i^{(k)}(\mathbf{B})$ is defined similarly. The set $\mathcal{Z}_i^{(k)}(\mathbf{A})$ is the set of assignments in which unit i and all its $k - step$ neighbors are assigned to A. Consider the following Horvitz-Thompson estimator for the average causal effect:

$$\begin{aligned}\hat{\theta}(\mathbf{Z}) &= \hat{\theta}_{\mathbf{A}}(\mathbf{Z}) - \hat{\theta}_{\mathbf{B}}(\mathbf{Z}) \\ &= \frac{1}{N}\sum_i \frac{I(Z \in \mathcal{Z}_i^{(k)}(\mathbf{A}))}{P(\mathbf{Z} \in \mathcal{Z}_i^{(k)}(\mathbf{A}))} Y_i(\mathbf{A}) \\ &\quad - \frac{1}{N}\sum_i \frac{I(Z \in \mathcal{Z}_i^{(k)}(\mathbf{B}))}{P(\mathbf{Z} \in \mathcal{Z}_i^{(k)}(\mathbf{B})),} Y_i(\mathbf{B})\end{aligned}$$

which can be shown to be unbiased under $k-local$ interference (Aronow and Samii forthcoming). Under the BD, the variance has the following expression:

Proposition 2.

$$\mathbb{V}_{\mathcal{R}}[\hat{\theta}] = \mathbb{V}_{\mathcal{R}}[\hat{\theta}_{\mathbf{A}}] + \mathbb{V}_{\mathcal{R}}[\hat{\theta}_{\mathbf{B}}] - 2\mathbb{C}ov_{\mathcal{R}}[\hat{\theta}_{\mathbf{A}}, \hat{\theta}_{\mathbf{B}}],$$

where

$$\mathbb{V}_{\mathcal{R}}[\hat{\theta}_{\mathbf{A}}] = \tfrac{1}{N^2}\left[\sum_i (2^{|\mathcal{N}_i^{(k)}|} - 1)Y_i(\mathbf{A})^2 + \sum_i \sum_{j \neq i} (2^{|\mathcal{N}_i^{(k)}(A) \cap \mathcal{N}_j^{(k)}(\mathbf{A})|} - 1)Y_i(\mathbf{A})Y_j(\mathbf{A})\right]$$

and similarly for $\mathbb{V}_{\mathcal{R}}[\hat{\theta}_{\mathbf{B}}]$, and

$$\mathbb{C}ov_{\mathcal{R}}[\hat{\theta}_{\mathbf{A}}, \hat{\theta}_{\mathbf{B}}] = -\tfrac{1}{N^2}\left[\sum_i Y_i(\mathbf{A})Y_i(\mathbf{B}) + \sum_i \sum_{j \neq i} Y_i(\mathbf{A})Y_j(\mathbf{B})I(|\mathcal{N}_i^{(k)} \cap \mathcal{N}_j^{(k)}| > 0)\right].$$

The variance of the unbiased estimator thus depends explicitly on network quantities. To make things even more explicit, the next example focuses on a specific family of networks and shows that whether the asymptotic variance of the estimator converges depends on a single parameter of the network family.

Example 3. For this example, we will assume that there exist two scalars $K, M < \infty$ such that

$$0 < K < Y_i(\mathbf{Z}) < M, \quad \forall i = 1, \ldots, N, \quad \forall \mathbf{Z}$$

and will focus on the case where $k = 1$. If we model the network as an Erdos-Renyi graph with probability p of connection between nodes, we can show that

$$\mathbb{E}_{\mathcal{G}}\left[\mathbb{V}_{\mathcal{R}}[\hat{\theta}]\right] = O\left(\frac{2(1+p)^{N-1} - 1}{N} + (1+3p)(1+p^2)^{(N-2)} - 1 + \frac{1}{N} + \frac{N(N-1)}{2}(1-(1-p)(1-p^2)^{N-2})\right).$$

The behavior of this quantity depends on the parameter p, which governs the sparsity of the network. We show in the supplemental online appendix that

- if $p < \frac{1}{N}$, we have: $\mathbb{E}_{\mathcal{G}}\left[\mathbb{V}_{\mathcal{R}}[\hat{\theta}] \leq O(1/N)\right.$,
- if $p \geq 1/\sqrt{N}$, we have: $\mathbb{E}_{\mathcal{G}}\left[\mathbb{V}_{\mathcal{R}}[\hat{\theta}] \geq \frac{4e^{\frac{N-1}{\sqrt{N}}}}{N} K^2\right.$.

In this case, the expected variance of the estimator goes to zero if $p < 1/N$ but not if $p \geq 1/\sqrt{N}$.

It is straightforward to construct unbiased estimators under most forms of localized interference by relying on the popular Horvitz-Thompson estimators (Aronow and Samii forthcoming). On the other hand, checking that the variance of such estimators converges—even under simple forms of interference—can be difficult. Yet, Example 3 shows the perils of neglecting this arduous task, especially in large samples.

Remark 2. Aronow and Samii (forthcoming) state sufficient conditions for the consistency of Horvitz-Thompson estimators under a form of nested-population asymptotics; these conditions put implicit constraints on the network. While this is a useful approach, we found it more illustrative to structure the network directly using a network model (e.g., Erdos-Renyi), as in Example 3. In the supplemental online appendix, we compare the two perspectives and establish some connections.

4. DISCUSSION

4.1. *Broader Class of Estimands*

The estimands of the form $\theta(\mathbf{A}, \mathbf{B})$, which we have considered in this article, are relevant whenever the purpose of the experiment is to decide which of treatment **A** or treatment **B** would be best if applied to the entire population. This is the kind of question social platforms care about when they experiment with new products or features (Eckles, Kizilcec, and Bakshy 2016; Gui et al. 2015). Other scenarios, however, would call for different kinds of estimands. It should be clear that the fundamental problems raised by arbitrary interference do not vanish when more complex estimands are considered, although results analogous to Section 2 do become harder to formulate, as we illustrate next. For the rest of this section, consider treatment A to be an "active treatment" while treatment B is a "control" or "no treatment" and define the *average primary causal effect*:

$$\theta = \frac{1}{N}\sum_i Y_i(Z_i = A, \mathbf{Z}_{-i} = \mathbf{B}).$$

We can state the equivalent of Proposition 1:

Proposition 3. Denote by $\mathbf{Z}^{(i)}$ the assignment such that $Z_i = A$ and $Z_j = B$ for all $j \neq i$. Under arbitrary interference, the only unbiased estimators of θ under BD are of the form

$$\hat{\theta}(\mathbf{Z}) = C(\mathbf{Z}) + \frac{2^N}{N}\sum_i I(\mathbf{Z} = \mathbf{Z}^{(i)}) Y_i(\mathbf{Z}^{(i)}), \tag{3}$$

where $\sum_{\mathbf{Z}} C(\mathbf{Z}) = 0$.

The statement of Proposition 3 is more complex, but the problems it highlights are identical. Similar lines of reasoning hold for more complex estimands under arbitrary interference.

4.2. *Effective Treatments and Informative Sets*

In Section 3, the outcome of unit i depends on the assignment of units in its k-step neighborhood $\mathcal{N}_i^{(k)}$. This concept can be generalized by defining the *reference group* of user i (Manski 2013) to be the smallest set of units $G_i \subset \{1, \ldots, N\}$ such that

$$Y_i(\mathbf{Z}) = Y(\mathbf{Z}_{G_i}) \quad \forall \mathbf{Z}. \tag{4}$$

This generalizes Equation 2 and suggests that Z_{G_i}, called the *effective treatment* of unit i (Aronow and Samii forthcoming; Manski 2013), is more relevant than its treatment Z_i. Under Assumption 2 (no interference), the treatment and effective treatment of unit i are the same $\mathbf{Z}_{G_i} = Z_i$, while under under Assumption 3, the effective treatment of unit i encompasses the assignment of its k-step neighbors $\mathbf{Z}_{G_i} = Z_{\mathcal{N}_i^{(k)}}$. Under arbitrary interference, the effective treatment of unit i is the entire assignment vector, $\mathbf{Z}_{G_i} = \mathbf{Z}$. In the language of Section 2, an assignment vector $\mathbf{Z}'$ will be informative for the outcome $Y_i(\mathbf{Z})$ if it results in the same effective treatment. We thus define the *informative set* for $Y_i(\mathbf{Z})$:

$$\mathcal{Z}_i(\mathbf{Z}) = \{\mathbf{Z}' : \mathbf{Z}'_{G_i} = \mathbf{Z}_{G_i} \quad \textit{and} \quad \mathbb{P}_{\mathcal{R}}(\mathbf{Z}') > 0\}.$$

Under Assumption 2, the informative set for any outcome $Y_i(\mathbf{Z})$ contains every assignments $\mathbf{Z}'$ such that $Z'_i = Z_i$, while under arbitrary interference, its informative set only contains the assignment $\mathbf{Z}$. The next section explores how the concepts of effective treatment and informative sets capture the important changes implied by different interference structures.

4.3. *Understanding Interference Structures*

Although we have defined effective treatments and informative sets for abstract designs $\mathcal{R}$, we focus the rest of the discussion on the BD, which captures the salient features of the problem while simplifying the exposition. Denote by $E_i = |\{\mathbf{Z}_{G_i}\}_{\mathbf{Z}}|$ the number of effective treatments for unit i and $S_i(\mathbf{Z}) = |\mathcal{Z}_i(\mathbf{Z})|$ the size of the informative set for the outcome $Y_i(\mathbf{Z})$. We call $F_i(\mathbf{Z}) = S_i(\mathbf{Z})/|\mathcal{Z}|$ the fraction of informative assignments for $Y_i(\mathbf{Z})$. Under the BD, we show in the supplemental online appendix that

$$\forall i, \mathbf{Z} \quad F_i(\mathbf{Z}) = F_i \quad \text{and} \quad E_i = \frac{1}{F_i}, \tag{5}$$

establishing a connection between the number of effective treatments and the fraction of informative assignments for all units i. Table 1 illustrates this relation for different interference structures. We see that stronger assumptions on the interference structure tend to reduce the number of effective treatments or, equivalently, increase the fraction of relevant sets. Applying these insights to Example 3, Table 2 contrasts the behaviors of $\mathbb{E}_{\mathcal{G}}[E_i]$ and $\mathbb{E}_{\mathcal{G}}[F_i(\mathbf{Z})]$ as $N \to \infty$ for sparser networks ($p = 1/N$) and denser networks ($p = 1/\sqrt{N}$). For sparser networks, the expected number of effective treatments converges while the fraction of informative assignments converges to a strictly positive number. In contrast, denser networks ($p = 1/\sqrt{N}$) lead to an infinite number of effective treatments in expectation, as $N \to \infty$. Under assumption 3, a unit's outcome depends on the the assignment of its neighbors, and so the fewer neighbors it has, the closest it is to the no-interference scenario.

Table 1. Expected Number of Effective Treatments and Fraction of Informative Assignments for Different Interference Structures, under the Bernoulli Design

Interference	No	1-local	Arbitrary
E_i	2	$2^{\lvert\mathcal{N}_i\rvert}$	2^N
$F_i(\mathbf{Z})$	$\frac{1}{2}$	$\frac{1}{2^{\lvert\mathcal{N}_i\rvert}}$	$\frac{1}{2^N}$

Table 2. Asymptotic Expected Number of Effective Treatments and Fraction of **Z**-exposed Sets for the Two Different Erdos-Renyi Specifications of Example 3, under 1-local Interference.

	$p=1/N$	$p=1/\sqrt{N}$
$\mathbb{E}_{\mathcal{G}}[E_i]$	2e	∞
$\mathbb{E}_{\mathcal{G}}[F_i(\mathbf{Z})]$	$\frac{1}{2e^{1/2}}$	0

This intuition is supported by noticing that the column of Table 2 corresponding to the sparser networks ($p=1/N$) is closer to the column of Table 1 corresponding to the no interference case, while the denser networks are closer to the arbitrary interference case. Although none of the observations we made describe sufficient conditions for the consistency of our estimator (see Aronow and Samii forthcoming), they provide an intuitive connection between its asymptotic variance and the number of effective treatments (or equivalently, the fraction of informative assignments).

ORCID iD

Edoardo M. Airoldi https://orcid.org/0000-0002-3512-0542

Notes

1. This excludes estimands that are constant, namely, that do not depend on $\mathbf{Y}(\mathbf{A})$ and $\mathbf{Y}(\mathbf{B})$.
2. This term may incorporate covariates, as with *model-assisted* estimators. In the absence of external information, we focus on the case $C(\mathbf{Z})=0$.
3. See supplemental online appendix for technical details.

References

Aral, Sinan, and Dylan Walker. 2012. "Identifying Influential and Susceptible Members of Social Networks." *Science* 337(6092):337–41.

Aronow, Peter, and Cyrus Samii. Forthcoming. "Estimating Average Causal Effects under General Interference, with Application to a Social Network Experiment." *The Annals of Applied Statistics*.

Basse, Guillaume, and Edoardo Airoldi. Forthcoming. "Optimal Design of Experiments in the Presence of Network-correlated Outcomes." *Biometrika*.

Basse, Guillaume, and Avi Feller. Forthcoming. "Analyzing Two-stage Experiments in the Presence of Interference." *Journal of the American Statistical Association*.

Basu, Debabrata. 2011. "An Essay on the Logical Foundations of Survey Sampling, Part One." Pp. 167–206 in *Selected Works of Debabrata Basu*, edited by A. DasGupta. New York: Springer.

Bowers, Jake, Bruce Desmarais, Mark Frederickson, Nahomi Ichino, Hsuan-Wei Lee, and Simi Wang. 2018. "Models, Methods and Network Topology: Experimental Design for the Study of Interference." *Social Networks* 54:196–208.

Bowers, Jake, Mark Fredrickson, and Peter Aronow. 2016. "A More Powerful Test Statistic for Reasoning about Interference between Units." *Political Analysis* 24(3): 395–403.

Coppock, Alexander, and Neelanjan Sircar. 2013. "Design of Field Experiments under Unknown Interference Structures." Working paper.

Eckles, Dean, René F Kizilcec, and Eytan Bakshy. 2016. "Estimating Peer Effects in Networks with Peer Encouragement Designs." *Proceedings of the National Academy of Sciences* 113(27):7316–22.

Gui, Huan, Ya Xu, Anmol Bhasin, and Jiawei Han. 2015. "Network A/B Testing: From Sampling to Estimation." Pp. 399–409 in *Proceedings of the 24th International Conference on World Wide Web*, Florence, Italy.

Hudgens, Michael, and Elizabeth Halloran. 2008. "Toward Causal Inference with Interference." *Journal of the American Statistical Association* 103(482):832–42.

Imbens, Guido, and Donald Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge, UK: Cambridge University Press.

Manski, Charles F. 2013. "Identification of Treatment Response with Social Interactions." *The Econometrics Journal* 16(1):S1–23.

Rubin, Donald. 1980. "Comment." *Journal of the American Statistical Association* 75(371):591–93.

Särndal, Carl-Erik, Bengt Swensson, and JanWretman. 2003. *Model Assisted Survey Sampling*. Cambridge: Springer.

Shalizi, Cosma Rohilla, and Andrew Thomas. 2011. "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies." *Sociological Methods & Research* 40(2):211–39.

Sinclair, Betsy, Margaret McConnell, and Donald Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56(4):1055–69.

Ugander, Johan, Brian Karrer, Lars Backstrom, and Jon Kleinberg. 2013. "Graph Cluster Randomization: Network Exposure to Multiple Universes." Pp. 329–37 in

Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Author Biographies

Edoardo M. Airoldi is the Millard E. Gladfelter Professor of Statistics and Data Science and the director of the Center for Data Science in the Fox School of Business at Temple University. He received a PhD in computer science from Carnegie Mellon University in 2007; was a postdoctoral fellow at Princeton University; founded and directed the Harvard Laboratory for Applied Statistics and Data Science until 2017; and held visiting positions at Microsoft Research, Massachusetts Institute of Technology, and Yale University. His current research focuses on statistical theory and methods for designing and analyzing experiments on large networks, and, more generally, on modeling and inferential issues that arise in analyses that leverage network data. His work has appeared in journals across statistics, computer science, and general science, including *Annals of Statistics*, *Journal of the American Statistical Association*, *Journal of Machine Learning Research*, *Proceedings of the National Academy of Sciences*, and *Nature*. He has received an Alfred Sloan Research Fellowship, a Shutzer Fellowship from the Radcliffe Institute of Advanced Studies, an NSF CAREER Award, an ONR Young Investigator Program Award, and several best paper and other research awards. He delivered a plenary talk at the National Academy of Sciences Sackler Colloquium on “Causal Inference and Big Data” in 2015, and he gave an IMS Medallion Lecture at the Joint Statistical Meetings in 2017.

Guillaume W. Basse is a postdoctoral fellow in the Department of Statistics at University of California-Berkeley. He received a PhD in statistics from Harvard University in 2018, and will join the faculty of the Management Science and Engineering and Statistics departments at Stanford University in 2019. His current research focuses on the design and analysis of experiments on networks. His work has appeared in journals across statistics and computer science, including *Journal of the American Statistical Association, Biometrika*, and *Artificial Intelligence and Statistics*. He was awarded the Google Fellowship in Statistics for North America in 2014.