16. Estimating causal effects in experiments with network spillovers

Stephanie Zonszein, P. M. Aronow, and Cyrus Samii

INTRODUCTION

In an experiment with units that are connected to each other through some kind of network, it is common that treatments assigned to a given unit have effects that ripple through the network to create spillover effects on other units. For example, a sales pitch made to an individual could have effects that spill over beyond that individual onto peers in a social network if the pitch is shared or if peers respond to any changes in the recipient individual's behavior (Aral & Walker, 2011).

In the causal inference literature, we use the term "interference" to describe such situations in which a unit's outcomes depend on treatments administered to other units (Cox, 1958). Classical methods for analyzing experiments and conducting causal inference typically begin with an assumption of "no interference" (Imbens & Rubin, 2015). But interference and spillovers are common in many applied settings, so the "no interference" assumption is not tenable. This means that we need to depart from classical approaches and define new methods that can account for interference.

This chapter describes current *design-based* methods for estimation and inference for causal effects in randomized experiments when network spillovers are present. It draws primarily on Aronow and Samii (2017) and Aronow et al. (2021); readers are referred to those texts for more technical treatment as well as references to many other contributions in this area. This chapter also serves as an introduction to the interference package for the R statistical computing environment (Zonszein et al., 2022). The package offers methods for analyzing causal effects in the presence of spillover effects.

Analysis of network experiments can be rooted in either design-based or model-based inferential frameworks (Aral, 2016). The appeal of design-based methods is that inferential properties (e.g., conditions for unbiased, consistent, or efficient estimation) can be defined in terms of experimental designs, which are under the control of the analyst. These inferential properties can be defined in ways that are robust to different outcome distributions. Using design-based methods, an analyst can figure out how to design experiments that ought to control bias and improve efficiency in a way that is robust to a variety of application areas. The design-based approach can be contrasted with the model-based or structural approach, which requires the analyst to specify a parameterized model for the outcome data-generating process. In these model-based approaches, the validity of inferences depends strongly on whether the model is specified correctly. The design-based approach is often understood as a way to derive inferential properties without having to make strong assumptions about the outcome data-generating process. Of course, as we explain here, under interference, one must make *some* assumptions about the nature of the outcome data-generating process—namely, it is essential to start with

assumptions about network connections and the extent to which spillovers might travel on the network (Basse & Airoldi, 2018). That said, the approach described herein does not require that we make any strong assumptions about outcome distributions, linearity, or effect homogeneity, among other restrictions that are typically necessary in model-based approaches.

We begin by sketching out a formal setting based on the "potential outcomes" framework, extending it to account for interference. We then demonstrate how one can combine information about an experimental design, network connections, and forms of potential indirect exposures to treatment to make inferences on spillover effects. We work through applied examples that illustrate how different network topologies affect inferences in the presence of spillover effects.

NETWORK EXPERIMENTAL SETTING

We imagine a situation where an analyst is planning an experiment on a finite population, U, of units indexed by $i = 1,...,N$. The units are connected to each other through a network that we represent as an $N \times N$ adjacency matrix, θ, with entries that equal, for example, 0 or 1 (and with 0's on the diagonal). A value of 1 in row i and column j in the adjacency matrix means that i's potential outcomes might depend on whether or not j is treated (e.g., they may be immediate peers or close to each other in the network). A value of 0 in row i and column j means that i's potential outcomes most certainly do not depend on j's treatment status (e.g., there are no channels through which j could influence i, even indirectly). The undirected network in Figure 16.1 would be represented by a symmetric adjacency matrix, with 1's representing edges between units and 0's representing the absence of edges.

The experiment is one in which units could be assigned to either a "treatment" or "control" condition. We use an indicator, $z_i \in \{0,1\}$, to denote whether unit i is assigned to treatment ($z_i = 1$) or control $z_i = 0$). The $N \times 1$ vector of treatments for all units is given by $\mathbf{z} = (z_1,...,z_N)'$. Given that treatments are binary, there are, in principle, 2^N different treatment assignment profiles that could be constructed. An *experimental design* assigns some probability, $p_{\mathbf{z}}$, to a treatment profile, $\mathbf{z}$. Under the *Bernoulli assignment* experimental design, treatments are assigned through independent Bernoulli draws (e.g., independent coin flips), in which case $p_{\mathbf{z}} > 0$ for any of the 2^N possible profiles, although $p_{\mathbf{z}}$ would be vanishingly small for some profiles (such as the profile in which $\mathbf{z}$ is a vector of only 0's). The *complete*

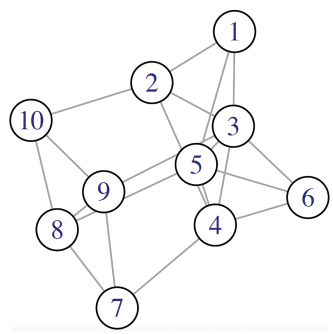

Figure 16.1 Example network

random assignment design, in which N_1 units are randomly assigned (without replacement) to treatment and $N_0 = N - N_1$ units to control, yields $p_{\mathbf{Z}} = 0$ for any assignment in which more or fewer than N_1 units are assigned to treatment, and $p_{\mathbf{Z}}$ would otherwise be uniform for all assignments with N_1 units assigned to treatment. Other types of experimental designs include various forms of stratified and clustered designs. Let the random $N \times 1$ vector $\mathbf{Z}$ refer to the treatment assignment profile that is ultimately realized, with $\Pr(\mathbf{Z} = \mathbf{z}) = p_{\mathbf{Z}}$.

We define causal effects in terms of how units' potential outcomes would change as $\mathbf{z}$ changes (Imbens & Rubin, 2015, ch. 1–2). We denote i's potential outcome under treatment assignment vector $\mathbf{z}$ as $y_i(\mathbf{z})$, where $y_i(\cdot)$ is a real-valued function that maps treatment profiles to an individual's outcome. For unit i, the effect of applying treatment profile $\mathbf{z}$ as compared to $\mathbf{z}'$ would be measured in the way that $y_i(\mathbf{z})$ differs from $y_i(\mathbf{z}')$. We focus on effects defined as differences, such as $y_i(\mathbf{z}) - y_i(\mathbf{z}')$. Let $\mathbf{z}_{-i} = (...,z_{i-1},z_{i+1},...)$ denote the $(N-1)$-element vector that removes the ith element from $\mathbf{z}$. Then we can also denote unit i's potential outcome as $y_i(z_i;\mathbf{z}_{-i})$. This allows us to characterize how effects for unit i might occur through direct exposure to the treatment (i.e., through variation in z_i, keeping $\mathbf{z}_{-i}$ fixed) or indirect spillovers (i.e., through variation in $\mathbf{z}_{-i}$, keeping z_i fixed). Traditional analyses of experiments assume no interference, which would imply that unit i's outcome depends only on the element z_i; that is, for any $\mathbf{z}$ and $\mathbf{z}'$, $y_i(z_i;\mathbf{z}) = y_i(z_i;\mathbf{z}')$. Interference means for some unit $i \in U$ we have $y_i(z_i;\mathbf{z}) \neq y_i(z_i;\mathbf{z}')$. The alternative extreme is when each value of $\mathbf{z}$ yields a different potential outcome for unit i, that is, for any $\mathbf{z} \neq \mathbf{z}'$, $y_i(\mathbf{z}) \neq y_i(\mathbf{z}')$. This would mean that unit i is subject to spillovers from any combination of other units.

EXPOSURE MAPPINGS AND CAUSAL EFFECTS

Following the approach of Aronow and Samii (2017), we define an *exposure mapping*, $f(\cdot)$, to establish what we think is the plausible extent and nature of spillover. For a unit i, the exposure mapping takes a treatment profile, $\mathbf{z}$, from the experimental design and the interference network, θ, as inputs, and then as an output, it generates an *exposure*, d. The exposure mapping is a way of modeling how different patterns of treatment assignment among peers might induce different potential outcomes. Thus, the exposure mapping translates potential outcomes defined in terms of treatment profiles, $\mathbf{z}$, into potential outcomes defined in terms of exposures, d. For example, we might suppose that a given unit's potential outcomes depend on whether (a) they are treated or (b) any one of their network peers is treated. This would imply an exposure mapping that takes as inputs the assignment profile $\mathbf{z}$ and unit i's row in the adjacency matrix, θ_i, as follows:

$$f(\mathbf{z},\theta_i) = \begin{cases} d_{11}(\text{Direct} + \text{IndirectExposure}): & z_i\mathbf{I}(\mathbf{z}'\theta_i > 0) = 1, \\ d_{10}(\text{IsolatedDirectExposure}): & z_i\mathbf{I}(\mathbf{z}'\theta_i = 0) = 1, \\ d_{01}(\text{IndirectExposure}): & (1-z_i)\mathbf{I}(\mathbf{z}'\theta_i > 0) = 1, \\ d_{00}(\text{NoExposure}): & (1-z_i)\mathbf{I}(\mathbf{z}'\theta_i = 0) = 1 \end{cases}$$

This particular exposure mapping implies that each unit i has four potential outcomes,

$$(y_i(d_{11}), y_i(d_{10}), y_i(d_{01}), y_i(d_{00})),$$

which are based on combinations of their own treatment status and whether any of their peers is treated. One might imagine other possible exposure mappings that suggest potential outcomes might depend on the number of peers treated, on whether peers of peers are treated, and possibly on other covariate values. The analyst is responsible for establishing a plausible specification. Then, analysis and inference proceed based on the specified exposure mapping. Aronow and Samii (2017, "Misspecification") explores the implications of misspecifying the exposure mapping, and Aronow et al. (2021, "Misspecified Exposure Mappings") demonstrates how bias and efficiency are influenced by specifying an exposure mapping that is either too coarse or too granular.

These exposure-based potential outcomes define causal effects. Under the exposure mapping defined above, the effect for unit i of being treated along with their peers, as compared to being in a situation where neither the unit nor any of their peers are treated, is given by $y_i(d_{11}) - y_i(d_{00})$. Insofar as this effect differs from the effect of isolated direct exposure ($y_i(d_{10}) - y_i(d_{00})$), then network spillover effects are meaningful in altering outcomes for those who are directly treated. Once the treatment profile, $\mathbf{Z}$, is drawn, units will be assigned to the corresponding exposure, d, and will reveal the associated potential outcome, $y_i(d)$. For a given unit, other exposures are counterfactual and not directly observable. For this reason, we do not focus on trying to estimate unit-level effects, but rather on average effects for the entire population,

$$\tau(d, d') = \frac{1}{N}\sum_{i=1}^{N}(y_i(d) - y_i(d')),$$

where d and d' are two different possible exposures.

DESIGN-BASED INFERENCE

Our goal is to obtain unbiased or consistent estimates of $\tau(d, d')$ effects and to be able to obtain informative and well-performing confidence intervals for statistical inference. With respect to bias, the issue in this setting is the nonrandom assignment of exposures *despite* random treatment assignment. Individual exposure is a function of both random treatment assignment and an individual's network position, captured by their adjacency matrix row, θ_i. Network position is nonrandom, and so this means that exposures are endogenous to potential confounding factors. For example, someone who has many direct peers has a much higher probability of receiving indirect exposure than someone with only a few peers. A unit's number of peers could be related to nontreatment factors that affect outcomes, introducing confounding. Based on classical results from sampling theory, Aronow and Samii (2017) showed that such confounding can be removed with inverse probability weighting—specifically, by weighting the outcome data contributed by a given unit by the inverse of the probability of the exposure received by that unit. Moreover, with the network adjacency matrix and exposure mapping in hand, these probabilities of exposure can be computed directly from the experimental design. Specifically, one can simulate treatment assignments based on the experimental design, feed them through the exposure mapping for each unit, and then determine each unit's probabilities of falling into any of the exposure conditions. The probabilities can be used to construct unbiased Horvitz-Thompson estimators of the $\tau(d, d')$ effects or consistent and considerably more efficient Hajek or weighted least squares estimators.

When it comes to statistical inference, the issues in this setting are the potentially complex dependencies across units, both in terms of outcomes but also, crucially, in the exposures that they receive. In the network in Figure 16.1, we can see that units 1 and 2 are both connected to unit 3, creating a dependency on whether these two units will receive indirect exposure. However, unit 2 is also tied to other peers of unit 3: units 5, 4, and 1. Such dependencies in exposure are a form of "multi-way clustering" that must be accounted for in statistical inference. Aronow and Samii (2017) propose sample theoretic methods for doing this in a way that conservatively accounts for variance components that cannot be directly estimated from the data.

EXAMPLES

We now turn to examples that put the foregoing ideas to work in estimating the direct and indirect effects of interventions under network spillovers. We do so using tools that are included in the interference R package. We use simulations to examine how differences in the network topology affect our inferences. In particular, how do differences in the skew in in-degree distributions affect inferences? Given that such skew introduces corresponding skew in the probabilities of indirect exposure, by the properties of inverse probability weighted estimators, we would expect that skew in degree contributes to higher variance and thus root mean square error (Robins et al., 2007). This is, in fact, what we see.

We work with the exposure mapping specified above, which assumes that experimental units are indirectly exposed to treatment when at least one first-degree neighbor is directly exposed to treatment (i.e., randomly assigned by the analyst to receive treatment). Given this assumption, we consider a scenario in which experimental units are exposed to one of four possible exposure conditions: direct + indirect exposure, isolated direct exposure, indirect exposure, and no exposure. Note that these exposure values result from considering the (direct) assignment of units into treatment and the topology of the network. Recall that a unit in the *direct + indirect exposure* is one that is (directly) assigned to treatment and also has a first-degree neighbor that is assigned to treatment. On the other end, units in the *no exposure* condition are not assigned to treatment and do not have first-degree neighbors assigned to treatment. Naturally, the likelihood of falling into the *isolated direct exposure*, for example, is higher for low-degree units (or with few first-degree neighbors) than for high-degree units (with many first-degree neighbors). We compute the exposure probabilities and estimate the exposure-specific causal effects using the inverse probability weighting methods described previously.

We consider two different network types. First, we construct a small-world network with $N = 1000$ units (vertices), an average degree 10, and a rewiring probability of 0.10. The small-world model of networks (Watts & Strogatz, 1998) is notable for featuring high clustering, like regular lattices, and a small diameter, like random graphs. Many social networks, including online social networks (Ugander et al., 2011), exhibit a high degree of clustering as well as a small diameter. Because the rewiring probability is closer to 0 than 1, in this case, we construct a small-world network that is closer to a regular lattice than a random graph. Most units are not neighbors of one another, but the neighbors of any given units are likely to be neighbors of each other, and most units can be reached from every other unit by a short path. Given such characteristics, the variation in vertex degree in small-world networks is limited.

The second network we construct is a degree-corrected stochastic block model (DCBM) (Karrer & Newman, 2011), featuring relatively high-clustering, but more variation in vertex

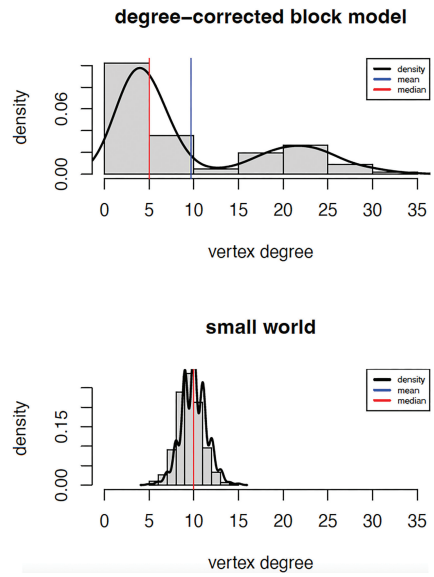

Figure 16.2 Comparison of vertex degree distribution

degree than is possible to produce with small-world networks. In particular, we generate a network according to a DCBM with $N = 1000$ units, 10 communities (or clusters); where units are assigned to one of the 10 communities uniformly at random, and an expected proportion of edges that is within a community of 0.7. The expected average degree in this network is 10 (as with the small-world network), and the variance is 75 (compared to a variance of 2 in the small-world network). Besides featuring substantially more variation in degree than the small-world network, the vertex degree distribution in this DCBM is also substantially more skewed (with a skewness of 1.07) than the vertex degree distribution in the small-world network (with a skewness of 0.12). Figure 16.2 illustrates the comparison in the distribution of vertex degrees across these two networks. We specifically choose these two networks to illustrate that the estimates of exposure-specific causal effects are more accurate and precise when the network exhibits less skewness in its vertex degree distribution.

To compute the exposure probabilities, we first need to simulate the set of all possible (or a large random sample of all possible) treatment assignments according to the experimental design. In this case, a proportion $p = 0.1$ of the $N = 1000$ units in the network is randomly assigned to treatment with uniform probability on the basis of complete random assignment. Because the set of all possible treatment assignments is very large ($\binom{N}{pN}$), we compute a set of R = 10,000 random treatment assignments. The R code for the `interference` package is as follows:

```
N <- 1000

p <- 0.1

possible_tr_assignment <- make_tr_vec_permutation(N, p, R
= 10000, seed = 4224)
```

We also need the adjacency matrix of the small-world and DCBM networks, as specified above,[1] which we compute with these functions,

```
make_adj_matrix_small_world <- function(N, seed) {
  set.seed(seed)
   g <- igraph::sample_smallworld(dim=1, size=N, nei=5,
p=0.1)
  while (min(igraph::degree(g, igraph::V(g))) == 0) {
    g <-
     igraph::sample_smallworld(dim=1, size=N, nei=5, p=0.1)
  }
  return(as.matrix(igraph::as_adj(g)))
}
adj_matrix_sw <- make_adj_matrix_small_world(N, seed = 42)
make_adj_matrix_dcsbm <- function(N, seed) {
  set.seed(seed)
  g <-
    randnet::BlockModel.Gen(10, N, K=10, rho = 0.7, alpha
= 1)
  while (min(rowSums(g$A)) == 0) {
    g <-
      randnet::BlockModel.Gen(10, N, K=10, rho = 0.7, alpha
= 1)
  }
  return(g$A)
}
adj_matrix_dcbm <- make_adj_matrix_dcsbm(N, seed = 42)
```

Next, we specify the exposure mapping function that transforms treatment assignments into exposure conditions, given the topology of the network and the type of assumed interference. Here, we assume that interference happens only through first-degree connections, and therefore we rely on an exposure mapping function that produces the four exposure conditions described previously (direct + indirect exposure, isolated direct exposure, indirect exposure, and no exposure). (Note that the package allows users to define and pass any function of the

network and treatment assignment vector.) With these three elements, we can now compute the exposure probabilities for each unit in each of the two networks.

```
prob_exposure_sw <- make_exposure_prob(possible_tr_
assignment,
                                       adj_matrix_sw,
                                      make_exposure_map_AS,
                                       list(hop = 1))
prob_exposure_dcbm    <-    make_exposure_prob(possible_tr_
assignment,
                                         adj_matrix_dcbm,
                                      make_exposure_map_AS,
                                         list(hop = 1))
```

These exposure probabilities are generalized propensity scores that can then be used in Horvitz–Thompson estimators or other inverse probability weighted estimators, such as the Hajek estimator of exposure-specific causal effects.[2] Compared to the Hajek estimator, the Horvitz–Thompson often suffers from excessive variance, and although the bias of the Hajek estimator is not 0, it is typically small and worth the variance reduction.

We now estimate exposure-specific effects. These are defined as contrasts between averages of individual potential outcomes across different exposures. To estimate the average potential outcome in a specific exposure condition, we have to take into account that we only observe the potential outcome for units in that specific exposure condition, and that the probability of observing potential outcomes in that specific exposure condition are not equal across units. This is precisely what inverse probability estimators like the Horvitz–Thompson take into account and the reason why we can obtain unbiased estimates of average potential outcomes in a specific exposure condition, and unbiased estimates of contrasts between these averages across different exposures.

To show how to compute exposure-specific effects, we simulate potential outcomes that exhibit effect heterogeneity and that vary in units' network degree. Specifically, we draw values for the baseline *no exposure* condition from an absolute standard normal distribution, which is correlated with the unit's first- and second-order degree. The values for the other exposure conditions are obtained by multiplying a vector of multipliers by the baseline *no exposure* value. The multipliers take values (2,1.5,1.25) for the *direct + indirect exposure*, *isolated direct exposure*, and *indirect exposure*, respectively, implying positive spillover effects.

```
potential_outcome_sw <- make_dilated_out(adj_matrix_sw,
                                         make_corr_out,
                                         hop = 1,
```

```
                                                    seed = 547)
potential_outcome_dcbm <- make_dilated_out(adj_matrix_dcbm,
                                           make_corr_out,
                                           hop = 1,
                                           seed = 547)
```

We also need to assign treatment and compute units' received exposure condition according to our assumed exposure scenario.

```
observed_tr_assignment <- make_tr_vec_permutation(N, p, R
= 1, seed = 4224)
received_exposure_sw <- make_exposure_map_AS(adj_matrix_sw,
                                     observed_tr_assignment,
                                              hop = 1)
received_exposure_dcbm <- make_exposure_map_AS(adj_matrix_
dcbm,
                                     observed_tr_assignment,
                                                hop = 1)
```

Finally, we get the observed outcomes from the potential outcomes and received exposures.

```
observed_outcome_sw <- rowSums(received_exposure_
sw*t(potential_outcome_sw))
observed_outcome_dcbm <- rowSums(received_exposure_
dcbm*t(potential_outcome_dcbm))
```

Next, we estimate exposure-specific effects using both the Horvitz–Thompson and Hajek estimators, as well as the variance estimators of these effects, which are conservative approximations to the exact variances. Particularly, we estimate exposure effects as a contrast to the *no exposure* condition (or the control condition). For example, the interactive effect of direct and indirect exposure results from the contrast between estimates of the average potential outcome in the *direct + indirect exposure* and the average potential outcome in the *no exposure* condition. The effect of direct exposure in the absence of any interaction with indirect exposure results from contrasting the estimates of the average potential outcome in the *isolated direct exposure* and the average potential outcome in the *no exposure* condition, and the effect of indirect exposure in the absence of any interaction with direct exposure from contrasting the average potential outcome in the *indirect exposure* and the average potential outcome in the *no exposure* condition.

```
estimates_sw <- estimates(received_exposure_sw,
                          observed_outcome_sw,
                          prob_exposure_sw,
                          control_condition = 'no')
estimates_sw$tau_ht
## dir_ind1 isol_dir     ind1
## 478.5162 174.5077 115.7963
estimates_sw$var_tau_ht
## dir_ind1 isol_dir       ind1
## 26198.147 52201.813 3038.021
estimates_sw$tau_h
## dir_ind1 isol_dir     ind1
## 456.9802 187.6371 123.1371
estimates_sw$var_tau_h
## dir_ind1 isol_dir       ind1
## 16088.628 17991.860 1017.804
estimates_dcbm <- estimates(received_exposure_dcbm,
                            observed_outcome_dcbm,
                            prob_exposure_dcbm,
                            control_condition = 'no')
estimates_dcbm$tau_ht
## dir_ind1 isol_dir     ind1
## 891.0871 664.6318 438.2607
estimates_dcbm$var_tau_ht
## dir_ind1 isol_dir       ind1
## 176098.04 816046.47 18307.43
estimates_dcbm$tau_h
## dir_ind1 isol_dir     ind1
## 766.9886 438.7036 291.4125
estimates_dcbm$var_tau_h
## dir_ind1 isol_dir       ind1
## 132124.84 318296.98 10881.17
```

Table 16.1 Estimates of Exposure Effects Across Two Networks

		Small-World			**Degree-Corrected Block Model**		
		(d_{11}, d_{00})	(d_{10}, d_{00})	(d_{01}, d_{00})	(d_{11}, d_{00})	(d_{10}, d_{00})	(d_{01}, d_{00})
*HT	$\hat{\tau}(d, d')$	478.51	174.50	115.79	891.08	664.63	438.26
	$s.e.(\hat{\tau})$	161.85	228.47	55.11	419.64	903.35	135.30
	$t_{\hat{\tau}(d,d')}$	2.95	0.76	2.10	2.12	0.73	3.23
*Hajek	$\hat{\tau}(d, d')$	456.98	187.63	123.13	766.98	438.70	291.41
	$s.e.(\hat{\tau})$	126.84	134.13	31.90	363.48	564.17	104.31
	$t_{\hat{\tau}(d,d')}$	3.60	1.39	3.85	2.11	0.77	2.79

Note: $\hat{\tau}(d, d')$ is effect estimate, $s.e.(\hat{\tau})$ is standard error, and $t_{\hat{\tau}(d,d')}$ is *t*-statistic.

Table 16.1 summarizes and compares these estimates across the two networks. We see that the *t*-statistic of the effect estimates are generally smaller for the DCBM, which exhibits substantial skewness in vertex degree distribution.

To talk more about the performance of the estimators across these two networks, we now present a simulation that replicates the experiment 5000 times by reassigning treatment each time. Our evaluation metrics are bias and root mean squared error (RMSE), which is increased by both bias and variance of the estimated exposure-specific effects.

Figure 16.3 presents the results of the simulation. Comparing the performance of the estimators across the two networks, we see that the bias and RMSE are bigger for the DCBM, which exhibits a larger vertex degree heterogeneity and substantial skewness in the vertex degree distribution. Comparing the performance of the two estimators, we see that with the Hajek estimator, the variance reduction is large and that it overwhelms the increase in bias such that Hajek reduces not only variance but also RMSE substantially; and that this is particularly visible for the DCBM with notable variation in vertex degree, which is the principal determinant of the propensities used by the Hajek estimator.

CONCLUSION

Classical approaches to causal inference and analysis of experiments define an effect on a unit solely in terms of the treatment that the unit itself receives. This "no interference" assumption is untenable in situations where treatments have effects that spill over from treatment recipients to peers in a network. Such spillover can occur through peer information sharing or reactions to peer behaviors that are affected by treatments.

This chapter presented a design-based approach for causal inference in network experiments with spillover effects. The approach requires combining information about the experimental design, the network, and then plausible ways through which spillovers might occur, although it does not require a parameterized specification of the outcome data-generating process. Reliable inference in such settings requires that one address (a) potential confounding based on

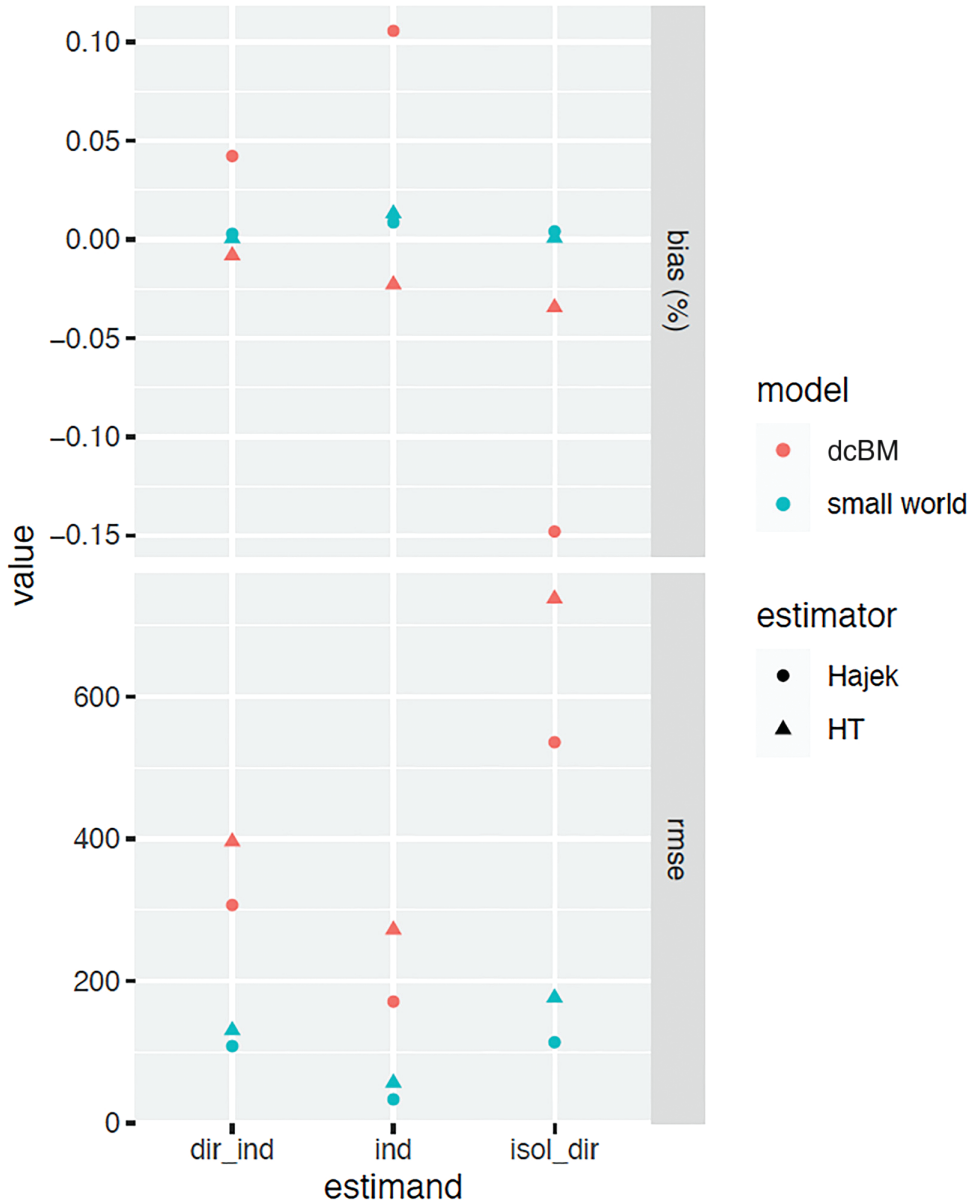

Figure 16.3 Estimator performance across networks

nonrandom variation in units' network locations and (b) complex dependencies in the ways that a set of units might be subject to spillover effects through shared peers. We described an approach that addresses these complications through inverse probability weighted estimators.

Network topology affects the precision of inferences. In a simulation study, we showed that a network with a high skew in the in-degree network (a degree-corrected stochastic block model) results in much noisier estimates than a network with a relatively tight degree distribution (namely, a small-world network). Thus, the degree distribution has important implications for the statistical power of network experiments. Analysts should construct similar simulations, encoding features of the networks that they plan to study in their experiments, in order to get assessments of statistical power.

FURTHER READING

Aral (2016), Taylor and Eckles (2018), and Aronow et al. (2021) describe many examples of network experiments in the social sciences and review a variety of methods for analyzing them. When the focus is on the average difference in unit potential outcomes between 100 percent and 0 percent treatment saturation, Ugander et al. (2013) and Eckles et al. (2017) suggest approaches for designing network experiments, specifically graph cluster randomization, to minimize the overall error in estimating the average effects of a global treatment. The interference package for R (Zonszein et al., 2022) includes functions for graph cluster randomization. Finally, Hudgens and Halloran (2008) propose an approach for analyzing network experiments when one can assume that interference is limited to occurring within well-defined and nonoverlapping groups, but the interference network per se is unknown. The interference package enables the application of such an approach.

NOTES

1. Our networks do not have isolates, but we suggest researchers remove isolates from their analysis.
2. For a formal discussion of the estimators, see Aronow and Samii (2017).

REFERENCES

Aral, S. (2016). Networked experiments. In Y. Bramoullé, A. Galeotti, & B. W. Rogers (Eds.),*The Oxford handbook of the economics of networks* (pp. 376–411). Oxford University Press.

Aral, S., & Walker, D. (2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*, *57*(9), 1623–1639.

Aronow, P. M., Eckles, D., Samii, C., & Zonszein, S. (2021). Spillover effects in experimental data. In J. N. Druckman & D. P. Green (Eds.), *Advances in experimental political science* (pp. 289–319). Cambridge University Press.

Aronow, P. M., & 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.

Basse, G. W., & Airoldi, E. M. (2018). Limitations of design-based causal inference and A/B testing under arbitrary and network interference. *Sociological Methodology*, *48*(1), 136–151.

Cox, D. R. (1958). *Planning of experiments*. Wiley.

Eckles, D., Karrer, B., & Ugander, J. (2017). Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, *5*(1), 20150021.

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

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

Karrer, B., & Newman, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, *83*(1), 016107.

Robins, J., Sued, M., Lei-Gomez, Q., & Rotnitzky, A. (2007). Comment: Performance of double-robust estimators when "inverse probability" weights are highly variable. *Statistical Science*, *22*(4), 544–559.

Taylor, S. J., & Eckles, D. (2018). Randomized experiments to detect and estimate social influence in networks. In S. Lehmann & Y.-Y. Ahn (Eds.), *Complex spreading phenomena in social systems* (pp. 289–322). Springer. doi: 10.1007/978-3-319-77332-2_16

Ugander, J., Karrer, B., Backstrom, L., & Kleinberg, J. 2013. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 329–337). ACM.

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the Facebook social graph. *arXiv preprint arXiv:1111.4503*.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, *393*(6684), 440–442.

Zonszein, S., Aronow, P. M., & Samii, C. (2022). *Interference, an R package for design-based estimation of spillover effects*. https://github.com/szonszein/interference.