Estimation of Causal Peer Influence Effects

Panos Toulis PTOULIS@FAS.HARVARD.EDU
Edward Kao EDWARDKAO@FAS.HARVARD.EDU
Department of Statistics, Harvard University, 1 Oxford Street, Cambridge, MA 02138, USA

Abstract

The broad adoption of social media has generated interest in leveraging peer influence for inducing desired user behavior. Quantifying the *causal effect* of peer influence presents technical challenges, however, including how to deal with social interference, complex response functions and network uncertainty. In this paper, we extend *potential outcomes* to allow for interference, we introduce well-defined causal estimands of peer-influence, and we develop two estimation procedures: a frequentist procedure relying on a sequential randomization design that requires knowledge of the network but operates under complicated response functions, and a Bayesian procedure which accounts for network uncertainty but relies on a linear response assumption to increase estimation precision. Our results show the advantages and disadvantages of the proposed methods in a number of situations.

1. Introduction

In causal inference, there is *interference* among units of analysis when a treatment or intervention on a unit has an effect on the response of another. These effects are known as *spillover effects* in economic theory, but when the units are humans and the effects originate from a unit's peers, e.g., friends or classmates, they are known as *peer influence effects*. The recent rise in adoption of social media has focused attention on how to quantify peer influence, and thereby raised technical challenges for causal inference when interference is present.

Recent works in statistics, for instance, account for peer influence (i.e., interference) as nuisance when estimating classical average treatment effects. In contrast, our work focuses on estimating the peer influence effect itself. Rosenbaum (Rosenbaum, 2007) takes a non-parametric approach, assuming only the existence of a baseline uniformity trial in which there may be interactions but there is no real treatment (placebo). To uncover the treatment effects under the "nuisance" of interference, one can compare the relative responses between treated and control units across groups or clusters of units. In (Hudgens & Halloran, 2008), Hudgens and Halloran consider the problem of estimating causal effects from vaccination under interference and identify these effects by comparing two reference groups: one group with low vaccination percentage and one group with a high percentage, under the key assumption of no interference among groups. A survey of the relatively thin literature on causal inference under interference is given by Tchetgen & VanderWeele (2012).

A related line of research in the social sciences aims at estimating peer influence effects in the absence of a true causal framework. Here, identifying whether a unit's outcome is a result of the social ties to peers (social contagion) or a result of similarity (homophily) to peers remains challenging (Manski, 1993). Applications of peer effects research have spanned many areas such as behavioral science/public health (Mednick et al., 2010), advertisement (Parker, 2011), network security (Shah & Zaman, 2011), and economics (Acemoglu et al., 2010). The effect of interference in these works is captured by a parameter in the model (e.g. in structural equation models), which nearly always lacks well-defined causal interpretation.

Applications for this problem abound. In (Bakshy et al., 2012), Facebook users were shown ads with and without their friends' product affiliation. The goal was to understand peer influence and, specifically, estimate the probability of sharing an "endorsement" conditioning on the strength of the friend ties. In (Bond et al., 2012) a massive randomized experiment on Facebook investigated peer influence effects on voting turnout. In a different setting, where peer effects can be "trans-

Proceedings of the 30^{th} International Conference on Machine Learning, Atlanta, Georgia, USA, 2013. JMLR: W&CP volume 28.

mitted" through an integrated market, (Ostrovsky & Schwarz, 2010) perform a large on-line randomized experiment to assess the impact of reserve prices in total revenue of Yahoo! ad auctions. Lacking a general methodological framework, most studies including the aforementioned ones, assume away interference for the sake of simplicity. The risk of this assumption is demonstrated in (Sobel, 2006). Examining housing mobility studies in which households in poor areas are financed to relocate to better neighborhoods, Sobel shows that ignoring interference can lead to entirely wrong conclusions about the effectiveness of the program.

Recently, the potential outcome framework is gaining traction in this problem space. Unpublished work by Ugander et al. (2012), for example, targets estimands related to ours with a different estimation procedure. Aronow & Samii (2013) perform inference using the Horwitz-Thompson estimator, assuming a network sampling design is in place and that the sampling inclusion probabilities can be computed. Our work differs in that (i) we propose a causal estimand that is well-defined and tightly-connected to the underlying network, (ii) we insist on sequential randomization designs which we believe are more appropriate for this problem and (iii) we consider a linear Bayesian model that can accommodate network uncertainty and increase precision under suitable conditions.

1.1. Causal frameworks & potential outcomes

The Rubin causal model (Rubin, 1974; 1990), based on potential outcomes, is the most widely-used causal framework in statistics and the social sciences. This approach is rooted in a fundamental question: How would we define a causal effect if we had all the data?

For the simplest scenario, assume two units, indexed by i, who are about to receive a treatment, say, an aspirin. Denote with $Z_i \in \{0, 1\}$ whether unit i received the treatment ($Z_i = 1$) or not ($Z_i = 0$) and $\mathbf{Z} = (Z_1, Z_2)$ the entire assignment vector. Also denote the response to the treatment, say, severity of headache, with $Y_i(Z_i)$ (see Table 1). The fundamental assumption that enables us to write $Y_i(Z_i)$ is that of *no interference*, also known as SUTVA (Stable Unit Treatment Value Assumption); that is, the outcome of individual i is only a function of its treatment Z_i.

Unit	"Ideal" world		Real world	
	$Z_i = 0$	$Z_i = 1$	$Z_i = 0$	$Z_i = 1$
1	$Y_1(0)$	$Y_1(1)$	$Y_1(0)$	?
2	$Y_2(0)$	$Y_2(1)$	?	$Y_2(1)$

Table 1. Causal inference as a missing data problem under the Rubin model

In an "ideal" world, we would observe all the possible outcomes (left part of table). In this ideal scenario, we define the *causal estimand* by pretending that we have access to outcomes for all possible treatment assignments. For example, a natural definition of the causal effect of taking aspirin would be:

$$\mu = (1/2) \cdot [Y_1(1) + Y_2(1) - Y_1(0) - Y_2(0)]$$

In the real world, however, only one outcome can be observed for each unit, because one cannot both take and not-take the aspirin, whereas the other will be missing (denoted with "?" in the table). A desirable feature of the potential outcomes framework is the ability to define causal estimands in terms of individual-level potential outcomes—even though only typical (e.g. average) causal estimands are estimable in practice.

Estimation proceeds in two ways. In randomization-based inference, treatment is randomized and estimates are obtained as functions of the observed outcomes. Here, if aspirin assignment was randomized, and unit 2 received aspirin, then a natural estimate would be $Y_2(1) - Y_1(0)$, and such an estimate would be unbiased[1]. In a model-based approach, the outcomes can be modeled conditioned on the assignment and the observed values e.g., assume $Y_i(Z_i)$ is normal with mean $\mu + \tau Z_i$. An alternative causal framework is that of causal graphical models (Pearl, 2000; Spirtes et al., 2001), which uses directed acyclic graphs (DAGs) to represent causal dependencies. This framework is popular in computer science, however, it is not well suited to our problem since we do not aim at estimating a causal structure but rather the "marginal" peer influence causal effects through a randomized experiment. Interestingly, identifiability of causal effects from observational data (even under fixed causal graphs) has recently been challenged (Shalizi & Thomas, 2011).

1.2. Contributions

This paper introduces a new and well-defined causal estimand for peer influence effects in Section 2.1, by extending potential outcomes to allow for interference in a social network. We then develop two ways to estimate this causal estimand, first through sequential randomization and second via a model-based ap-

[1] In general, the estimator $\hat{\mu} = Z_1 \cdot Y_1(1) + Z_2 \cdot Y_2(1) - (1 - Z_1) \cdot Y_1(0) - (1 - Z_2) \cdot Y_2(0)$ is unbiased since $E[Z_i] = 1/2$.

proach, and then demonstrate their trade-offs. Section 2.2 describes the randomization approach and characterizes the subtlety of this problem by introducing the idea of *manipulability* of a network. Sections 3.1 and 4.1 characterize the trade-off between manipulability and possible bias of a randomized design. Theorem 1 shows that this bias is intertwined with network-specific properties (e.g. the *sharing index*). Section 2.3 describes the model-based approach under a specific additivity assumption. Section 3.2 shows how to optimize asymptotic expected performance through assigning the treatment vector to maximize Fisher information, thus providing insights for experimental design.

2. Methods

We denote a network as $G = (V, E)$, where V is the vertex set, $|V| = N$, and E is the edge set. For a node $i \in V$, we define as $\mathcal{N}_i$ to be its neighborhood, excluding i. Node i has $n_i = |\mathcal{N}_i|$ neighboring nodes. The $N \times 1$ treatment assignment vector is denoted by $\mathbf{Z}$, where $Z_i = 0$ or 1 if node i is assigned to control or treatment, and for a subset $S \subseteq V$, let $\mathbf{Z}_S$ be the assignment vector for the nodes in S. Thus, $\mathbf{Z}_{\mathcal{N}_i}$ is the assignment vector of the neighbors of i. Also, let V_k be the set of nodes that have at least k neighbors and $\mathcal{M}_{ik}$ be the set of neighbors of node $i \in V_k$ who are also neighbors to at least one other node in V_k. Define $\mathcal{M}_k = \bigcup_i \mathcal{M}_{ik}$ as the set of *shared neighbors*. Denote also $m_{ik} = |\mathcal{M}_{ik}|$ i.e., the # of neighbors of i who are shared with other nodes in V_k as well.

We say that a node, or equivalently an experimental unit i is *treated* if $Z_i = 1$ and it is in control if $Z_i = 0$. When $Z_i = 1$, unit i is said to have *primary effects*. Furthermore, we say that a unit i is *exposed to peer influence effects* if at least one neighbor is treated. A unit $i \in V_k$ is *k-exposed* if exactly k neighbors are being treated and the corresponding treatment assignment is called a *k-level assignment*. We say that a unit is *non-exposed* when $Z_i = 0$ and $\mathbf{Z}_{\mathcal{N}_i} = \mathbf{0}$ i.e., the unit i and all its neighbors are in control. Also, denote with $\mathcal{D}_i$ the set of all assignments $\mathbf{Z}_{\mathcal{N}_i}$ that make node i to be k-exposed.

The response of unit i (potential outcome) under treatment $\mathbf{Z}$ is denoted by $Y_i(\mathbf{Z}) \equiv Y_i(Z_i, \mathbf{Z}_{-i})$, where $\mathbf{Z}_{-i}$ is the vector of assignment $\mathbf{Z}$ excluding i's assignment. Define $\mathbf{Z}(\mathcal{N}_i; k)$ to be the set of all assignments on $\mathcal{N}_i$ in which exactly k neighbors of i get treated (total $\binom{n_i}{k}$ such assignments). Define $\mathbf{Z}_1(\mathcal{N}_i; k)$ as the set of all assignments in $\mathbf{Z}(\mathcal{N}_i; k)$ for which $\exists j, Z_j = 1$ and $j \in \mathcal{M}_{ik}$, i.e., at least one of the shared neighbors of i gets treated. Denote as $\mathbf{Z}_0(\mathcal{N}_i; k)$ the set of assignments $\mathbf{Z}(\mathcal{N}_i; k) \setminus \mathbf{Z}_1(\mathcal{N}_i; k)$, i.e. node i is k-level exposed and all shared neighbors are put in control. Notice that it holds, $\mathbf{Z}_0(\mathcal{N}_i; k) \bigcup \mathbf{Z}_1(\mathcal{N}_i; k) = \mathbf{Z}(\mathcal{N}_i; k)$ and $\mathbf{Z}_0(\mathcal{N}_i; k) \bigcap \mathbf{Z}_1(\mathcal{N}_i; k) = \emptyset$, i.e. the two sets are disjoint and form collectively the entire set of k-level assignments.

Last, define $\rho_i = \binom{n_i}{k}^{-1}$, $\rho_{0,i} = \binom{n_i - m_{ik}}{k}^{-1}$ and $\rho_{1,i} = (\frac{1}{\rho_i} - \frac{1}{\rho_{0,i}})^{-1}$. Intuitively ρ_i is the probability of one random k-level assignment for unit i and $\rho_{0,i}$ is the probability of a random assignment *given* that i's shared neighbors are put in control. Note that $|\mathbf{Z}_0(\mathcal{N}_i; k)| = 1/\rho_{0,i}$ and $|\mathbf{Z}_1(\mathcal{N}_i; k)| = 1/\rho_{1,i}$, so that $\frac{1}{\rho_i} = \frac{1}{\rho_{0,i}} + \frac{1}{\rho_{1,i}}$.

2.1. Causal estimands for treatment effects

In the classical potential outcomes framework, SUTVA is assumed: $Y_i(\boldsymbol{Z}) = Y_i(Z_i)$, meaning that the outcome of unit i depends only on the treatment it receives and not on the treatment other units receive. This is clearly violated in the presence of interference. We replace SUTVA with the following, more relaxed, assumption:

Assumption 1. We assume $\forall i$, $Y_i(\mathbf{Z}) = Y_i(Z_i, \mathbf{Z}_{\mathcal{N}_i})$, i.e., a unit's response can be affected by the treatment it receives and by the treatments received by its neighbors. □

Formally, the response function of a node i can be denoted by $Y(Z_i, \mathbf{Z}_{\mathcal{N}_i})$ and is a map $\{0,1\} \times \{0,1\}^{|\mathcal{N}_i|} \to D_y$, where D_y is the domain of potential outcomes. For brevity, we denote the response $Y_i(0, \mathbf{Z}_{\mathcal{N}_i} = \mathbf{0})$ by $Y_i(\mathbf{0})$.

Definition 1. [Estimand for primary effects] Define as ξ the causal estimand of primary effects as follows.

$$\xi \equiv \frac{1}{N} \sum_i Y_i(1, \mathbf{Z}_{\mathcal{N}_i} = \mathbf{0}) - Y_i(\mathbf{0}) \tag{1}$$

Definition 2. [Main estimand for peer influence effects] Define as δ_k the causal estimand of k-level effects as follows:

$$\delta_k \equiv \frac{1}{|V_k|} \sum_{i \in V_k} \left[\binom{n_i}{k}^{-1} \sum_{\mathbf{z} \in \mathbf{Z}(\mathcal{N}_i; k)} Y_i(0, \mathbf{z}) - Y_i(\mathbf{0}) \right] \tag{2}$$

Definition 3. [Additional peer influence effects estimands]

("insulated neighbors")

$$\delta_{k,0} = \frac{1}{N} \sum_i \rho_{0,i} \cdot \sum_{\mathbf{z} \in \mathbf{Z}_0(\mathcal{N}_i; k)} Y_i(0, \mathbf{z}) - Y_i(\mathbf{0}) \tag{3}$$

("non-insulated neighbors"):

$$\delta_{k,1} = \frac{1}{N}\sum_i \rho_{1,i} \cdot \sum_{\mathbf{z}\in\mathbf{Z}_1(\mathcal{N}_i;k)} Y_i(0,\mathbf{z}) - Y_i(\mathbf{0}) \quad (4)$$

In the following sections, we describe methods to estimate our causal estimands. One key concept is that of the *valid* causal estimate, which is a measure of *treatment balance*:

Definition 4. [Valid causal estimates] A causal estimate from a randomization is *valid* if at least one node was assigned to the prescribed treatment and at least one node was assigned to control. Otherwise, the estimate is not valid.

Estimation of ξ is straightforward through a typical randomized experiment. However, estimation of δ_k is more involved[2] because of interference. In the following sections, we will focus on estimation of δ_k. Note that, by the definition of δ_k, any randomization needs to set $Z_i = 0$ for nodes $i \in V_k$ and randomize treatment only within their neighborhoods.

2.2. Causal inference through randomization

We start with a simple sequential design:[3]

Algorithm 1 Estimation of δ_k: Simple Sequential Randomization $\texttt{SSR}(G, \mathbf{Z})$
Input: G network, $\mathbf{Z}$ current treatment vector
Output: $\mathbf{Z}$ treatment vector (in-place)

1: **while** $i \leftarrow \texttt{sample}\{i : i \in V_k \ \&\ \texttt{s}(\mathbf{Z}_{\mathcal{N}_i}) \leq k\}$ **do**
2: $\quad T_i = \{j \in \mathcal{N}_i : Z_j \neq \text{NA}\}$
3: $\quad \mathbf{W} \leftarrow \texttt{sample}\{\mathbf{W} : \mathbf{W} \in \mathcal{D}_i \ \&\ \mathbf{W}_{T_i} = \mathbf{Z}_{T_i}\}$
4: $\quad \mathbf{Z}_{\mathcal{N}_i} \leftarrow \texttt{sample}\{\mathbf{W}, \mathbf{0}\}$
5: $\quad Z_i \leftarrow 0$
6: $\quad V_k \leftarrow V_k \setminus \{i\}$
7: **end while**

The SSR algorithm assigns nodes in V_k to a non-exposure or k-level exposure status. Specifically, in Line 3, a k-level assignment is sampled among those that maintain the treatment status of units who have already been assigned treatment. In Line 4, either a k-level assignment or non-exposure is finally chosen at random. Intuitively, SSR extends sequential randomization by taking into account the constraints of δ_k. However, the algorithm may come up with estimates that are not valid (see Definition 4). To illustrate, we refer to the "candy" network in Figure 1 (right). There are two nodes in V_k (orange nodes). Clearly, we can only get causal estimates when one node is k-exposed, the other is non-exposed, and we compare between the two observed outcomes. However, this can only happen when all middle nodes (shared neighbors) are put in control. The probability of this happening through SSR is very small[4].

A simple design that can alleviate this problem is presented in Algorithm 2. The randomization is essentially the same as SSR, but as a first step it puts $x\%$ of shared neighbors into control:

Algorithm 2 Estimation of δ_k: Insulated Neighbors Randomization $\text{INR}^x(G)$
Input: G network
Output: $\mathbf{Z}$ treatment vector

1: $Z_i \leftarrow \text{NA}, \forall i$
2: $S \leftarrow \texttt{sample}\{\texttt{n=}x \cdot |\mathcal{M}_k|, \mathcal{M}_k\}$
3: $\mathbf{Z}_S \leftarrow \mathbf{0}$
4: $\texttt{SSR(G,Z)}$

The idea behind INR is to increase the "causal information" acquired by a randomization at the expense of increased bias. We believe that this trade-off is key to estimating peer effects and we discuss more in Sections 3 and 4.

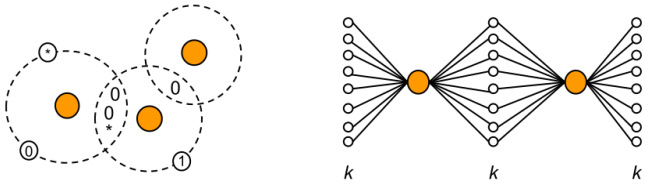

Figure 1. (Orange nodes are in V_k, "*" means unassigned) Left: $\texttt{INR}^x$ puts in control proportion of the common neighbors. Right: Candy network. To observe a causal estimate, the middle nodes need to be put in control. Only 2 out of all possible estimates are well-defined causal estimates.

[2]When defining our estimands, we could compare medians and not averages, ratios and not differences, or having a proportion of neighbors treated and not an exact number k. However, we believe our current approach to be conceptually clear and a good entry point to the problem.

[3]We assume function $\texttt{s(Z)}$ which counts how many units are treated in $\mathbf{Z}$ i.e., $\texttt{s(Z)} = |\{i : Z_i = 1\}|$. Function $\texttt{sample()}$ samples at random from a set and if an argument n is supplied then exactly n elements are sampled without replacement. Last, we assume that the value of Z_i for a unit i that has not been assigned treatment is equal to "NA" ("not assigned").

[4]Furthermore, since only 2 out of all causal estimates can be observed (at least $\binom{2k}{k}$ in total), any estimate will be severely biased, and the resulting estimator might have infinite risk.

2.3. Causal inference through linear model

The randomization based procedure above has the advantage of making no assumption about the node response functions $Y_i(.)$. However, it assumes complete knowledge of the network. Furthermore, depending on the network topology, it may have trouble finding enough valid causal estimates. Moreover, treatment assignments may be difficult to administer in real world scenarios.

The complementary model-based procedure addresses these issues by adopting a linear network treatment model that assumes additivity of the primary effects and peer influence effects in their contribution to the treatment response mean. The additivity assumption of various effects is also made by previous works, such as Manski's "linear in means" model (Manski, 1993). Our model is inspired by and similar to Parker's (Parker, 2011). As edges in real world networks often represent an uncertain quantity of interactions between two individuals, we extend it by considering weighted random networks.

The linear model assumes, in general, a weighted, undirected network G among units, with adjacency matrix A. We model the individual potential outcomes as:

$$Y_i(\boldsymbol{Z}) = \tau Z_i + \gamma \cdot \boldsymbol{a}_i'\mathbf{Z} + \mu + \varepsilon_i \tag{5}$$

in which $\boldsymbol{a}_i$ is the i-th column vector of the adjacency matrix A (in-links to unit i). This can be written in compact form as:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{6}$$

where $\boldsymbol{y} = (Y_i)$, the $N \times 1$ vector of responses, $\boldsymbol{X}$ is the $N \times 3$ design matrix such that $\boldsymbol{X} = [\mathbf{Z}, A'\mathbf{Z}, \mathbf{1}]$, $\boldsymbol{\beta} = (\tau, \gamma, \mu)'$ is the 3×1 parameter vector and the $N \times 1$ vector $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{I})$ is iid noise. Note that the network affects the likelihood only through the quantity $A'\mathbf{Z}$ i.e., the quantity $\boldsymbol{S} = A'\mathbf{Z}$ is the network sufficient statistic with respect to $\boldsymbol{\beta}$. Thus, $\boldsymbol{S}$ represents the amount of exposure to peer influence for each node.

2.3.1. Causal estimands under linear model

Under the linear model, the causal estimands are simplified and this helps bridge our two estimation procedures:

$$\begin{aligned}\xi &= \frac{1}{N}\sum_i \Big[Y_i(1, \mathbf{z} = \mathbf{0}) - Y_i(\mathbf{0})\Big] \\ &= \frac{1}{N}\sum_i \tau = \tau\end{aligned} \tag{7}$$

The k-level peer influence effects estimand δ_k reduces to a scaled γ:

$$\begin{aligned}\delta_k &= \frac{1}{|V_k|}\sum_i \left[\binom{n_i}{k}^{-1} \sum_{\mathbf{z}\in\mathbf{Z}(\mathcal{N}_i;k)} (Y_i(0,\mathbf{z}) - Y_i(\mathbf{0}))\right] \\ &= \frac{1}{|V_k|}\sum_i \left[\binom{n_i}{k}^{-1} \sum_{\mathbf{z}\in\mathbf{Z}(\mathcal{N}_i;k)} S_i(\mathbf{z})\gamma\right] \\ &= \frac{k\gamma}{|V_k|}\sum_i W_i \;\propto\; \gamma\end{aligned} \tag{8}$$

where W_i is the average weight on the incoming edges to unit i. Thus, estimating the causal estimands under the linear model amounts to inferring τ and γ.

2.3.2. Modeling network uncertainty

Real world networks are often uncertain, as true interactions between individuals may be either unobservable, or measured and estimated with error (Butts, 2003). Inspired by Perry's model for interaction networks (Perry & Wolfe, 2010), we model each edge weight (i, j) as a Poisson distributed random variable with rate λ_{ij}[5].

A key idea here is that, while the network G is random, we need to impute only its sufficient statistic $\boldsymbol{S}$, and thus inference can be efficient. In particular, we model $\boldsymbol{S}$ as follows:

$$A_{ij} \sim \text{Poisson}(\lambda_{ij}) \tag{9}$$

$$S_i = \boldsymbol{a}_i'\mathbf{Z} \sim \text{Poisson}(\kappa_i) \tag{10}$$

where $\boldsymbol{\kappa} = \boldsymbol{\Lambda}\mathbf{Z}$, and $\boldsymbol{\Lambda}$ is the $N \times N$ "interaction rate" matrix. Assuming we know the treatment assignment, the interaction rates and the unit responses, we arrive at the Bayesian model depicted in Figure 2.

Naturally we propose a joint inference procedure by treating the sufficient statistic $\boldsymbol{S}$ as the "missing data", and performing inference iteratively through MCMC with Gibbs sampling as shown in Figure 3. More details on this inferrential step are available in the supplementary material.

3. Theoretical results

3.1. Randomization performance analysis

It was argued in Section 2.2 that the network topology is important in getting causal estimates for δ_k. As another extreme example, consider the case of a complete

[5]For simplicity, here we assume knowledge of the rates. Future work may implement a more general hierarchical model on the rates as a function of some network parameters (e.g. node degrees in Chung-Lu model (William et al., 2001)).

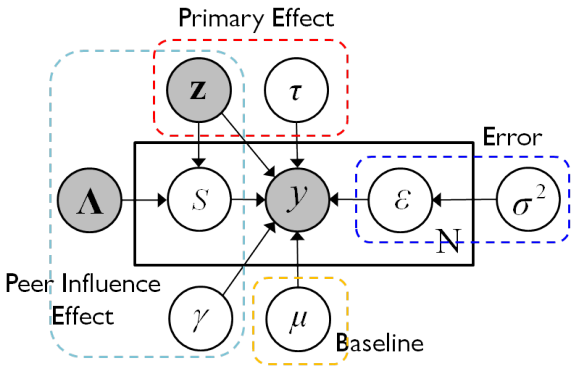

Figure 2. Plate diagram of the Poisson random network linear treatment model. $\boldsymbol{y}$, $\boldsymbol{\Lambda}$, and $\mathbf{Z}$ are known or observed.

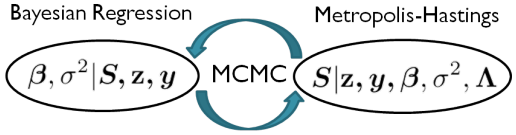

Figure 3. Joint inference of $\boldsymbol{\beta}$, σ^2, and $\boldsymbol{S}$ through MCMC.

graph (all units connected). Clearly, no causal estimate can be drawn from such a graph since as soon as a unit gets k-exposed no other unit can be non-exposed. Equivalently, for networks with isolated nodes, it is always possible to draw causal estimates through a completely randomized experiment. In other words, we can think of networks as having varying degrees of *manipulability*. We formalize this notion by the following definition:

Definition 5. [Manipulability] The *manipulability* of a network G under randomization R is the average proportion % of valid causal estimates that R will get when ran on G. □

Referring back to the previous examples, the complete graph has 0% manipulability and the isolated graph has 100% manipulability. Note that under our Assumption 1, interference happens through the shared neighborhood $\mathcal{M}_k$. Intuitively, the bigger $\mathcal{M}_k$ is with respect to the entire graph G the stronger the interference effects should be[6]. We formalize this notion by defining the *sharing index* of a network.

Definition 6. [Sharing index] For a given network and given k, the sharing index $\alpha \in [0, 1]$ is defined by:

$$\alpha = \frac{1}{|V_k|} \sum_{i \in V_k} \frac{m_{ik}}{n_i} \tag{11}$$

Intuitively, a high sharing index means that whenever a node is k-exposed or non-exposed, it will affect more of the remaining nodes, making the graph less manipulable[7]. Interestingly, the bias of the estimate provided by $\text{INR}^{1.0}$ is exactly proportional to the sharing index:

Theorem 1. If $\forall i, \rho_i = \rho$ and $\rho_{0,i} = \rho_0$, such that $\rho/\rho_0 = 1 - \alpha \leq 1$, then it holds:

$$E[\hat{\delta}_{k,INR}] = \delta_k + \alpha \cdot (\delta_{k,0} - \delta_{k,1}) \tag{12}$$

The bias of $\text{INR}^{1.0}$ estimates come from two sources: The first is from the sharing index of the network. The second comes from the difference of the influence that nodes in $\mathcal{M}_k$ (shared neighbors) exert compared with nodes in $V \setminus \mathcal{M}_k$. If for some reason (e.g. because of their better positioning on the network) the shared neighbors are more (or less) influential, our estimates will become biased. The following corollary summarizes these observations:

Corollary 1. *For an ego-centric network with no commonly shared nodes ($\alpha = 0$), the estimate from INR is unbiased. Furthermore, if peer influence effects are invariant to permutations of node ids (and so $\delta_{k,0} = \delta_{k,1}$) the estimate from INR is unbiased.*

3.2. Linear model performance analysis

Other than bias, the performance of the inference procedure is typically characterized by the variance of the estimate. Here we show how the structure of the network and treatment assignment play a role in the variance through the Fisher information matrix and the Cramér-Rao bound. From our Bayesian regression model, the likelihood is:

$$L_{\mathbf{y}}(\boldsymbol{\beta}, \boldsymbol{S}, \sigma^2 | \boldsymbol{\kappa}, \mathbf{Z}) = \prod_i \mathcal{N}(Y_i; \tau Z_i + \gamma S_i + \mu, \sigma^2) \times \text{Poisson}(S_i; \kappa_i) \tag{13}$$

To ease our calculations, we approximate the Poisson through a normal i.e., $\text{Poisson}(\kappa_i) \approx \mathcal{N}(\kappa_i, \kappa_i)$ and integrating out the nuisance parameter $\boldsymbol{S}$ by the normal representation, we take the log likelihood and compute the Fisher information matrix:

[6] For example, if $\mathcal{M}_k = \oslash$ then there is no interference and a classical randomized experiment can be performed.

[7] As a toy example consider a $G_{n,p}$ Erdös-Rényi graph. Then one can show that manipulability decreases with an exponential rate as the sharing index increases. However, this relationship is not yet fully characterized and can have a more complicated pattern in other types of networks.

$$\mathbf{I}(\boldsymbol{\beta}, \sigma^2) = -E\left(\frac{\partial^2 l_{\boldsymbol{y}}(\boldsymbol{\beta}, \sigma^2|\boldsymbol{\kappa}, \mathbf{Z})}{\partial[\tau, \gamma, \mu, \sigma^2]\partial[\tau, \gamma, \mu, \sigma^2]}\right)$$

$$= \begin{bmatrix} \sum_i \frac{z_i}{\phi_i} & \sum_i \frac{z_i\kappa_i}{\phi_i} & \sum_i \frac{z_i}{\phi_i} & 0 \\ . & \sum_i \left[\frac{\kappa_i^2}{\phi_i} - \frac{2\gamma^2\kappa_i^2}{\phi_i^2}\right] & \sum_i \frac{\kappa_i}{\phi_i} & 0 \\ . & . & \sum_i \frac{1}{\phi_i} & 0 \\ . & . & . & \sum_i \frac{1}{2\phi_i^2} \end{bmatrix} \tag{14}$$

where $\phi_i = \sigma^2 + \gamma^2\kappa_i$. The diagonal entries of the matrix reveal the information source for each parameter. Being mostly interested in the causal estimands τ and γ, we focus on I(1,1) and I(2,2). Since the rate of peer influence exposures $\boldsymbol{\kappa}$ are relatively large, the first term in I(2,2) dominates the second term. So we focus on $\sum_i \frac{z_i}{\sigma^2+\gamma^2\kappa_i}$ for I(1,1) and $\sum_i \frac{\kappa_i^2}{(\sigma^2+\gamma^2\kappa_i)}$ for I(2,2). Not surprisingly, the information content for τ comes from the treated nodes but is discounted by the rate of peer influence it receives. On the other hand, the information content for γ comes from having large amount of peer influences. This presents a tug of war between minimizing the variance on $\hat{\tau}$ and $\hat{\gamma}$. Interestingly, trying to meet both objectives leads to a treatment assignment where the hubs are treated to maximize overall peer influence while controlling their neighbors to minimize the peer influence to the treated nodes. These "isolated" treated hubs results in a treatment assignment strategy that is very similar to the INR randomization!

While the Cramér-Rao bound ($\boldsymbol{I}^{-1}$) is complicated and does not readily render intuition, one can compute it numerically and find the expected estimation variance. This can be used to numerically determine the minimal variance treatment assignment. Later we demonstrate empirically a reduction in estimation variance by using "optimal" treatment assignments (LMO) instead of random assignments (LMR).

4. Empirical results

4.1. Manipulability

The main idea behind INR is to increase manipulability of a network at the cost of introducing more bias. Consider, for example, the candy network on Figure 1. Under SSR, there is at most $2/\binom{2k}{k} \propto 4^{-k}$ probability of getting a valid causal estimate. In contrast, $\text{INR}^{1.0}$ will get causal estimates in 50% of the randomizations since all middle nodes (shared neighbors) will be put under control right away. As a further illustration, we test on Zachary's karate club network (Zachary, 1977). The summaries of SSR and $\text{INR}^{1.0}$ (in 10,000 samples, $k = 5, \delta_k = 2.542$) are shown below:

Design	Point Estimate	Manipulability	se
SSR	2.516	66.5%	1.022
$\text{INR}^{1.0}$	2.495	71.5%	1.062

Table 2. SSR vs. $\text{INR}^{1.0}$ randomizations and bias-causal information trade-off: INR introduces more bias but achieves more valid causal estimates.

$\text{INR}^{1.0}$, by fixing nodes under control, introduces more bias (first column) and more varied estimates. On the other hand, it obtains 5% more valid causal estimates on average than SSR thus increasing the manipulability of a rather dense social network.

4.2. Randomization vs model-based method

To introduce variety, we tested with 8 100-node networks, each with different type of topologies commonly seen in real-world networks. Two underlying response functions to the treatment are tested. The first one (Table 3) is based on the proposed linear response model $Y_i = \mu + \tau Z_i + \gamma S_i + \varepsilon_i$ and the second one (Table 4) adds a quadratic term ηS_i^2 to the response, signifying the phenomenon that somehow the peer influences reinforce each other. We pick the level-4 peer influence effects causal estimand, δ_4, as the objective. The true values for each response parameter were set to $\tau = 10$, $\gamma = 0.5$, $\mu = 3$, $\varepsilon_i \sim \mathcal{N}(0, 1)$, $\eta = 0.05$. The results for two levels of INR[8], the linear model with 15 random treatments (LMR) and 15 optimal treatments (LMO), are summarized in Tables 3 and 4.

Table 3 shows that when the model assumption is correct, the model-based approach out-performs the randomization-based approach, both in estimate bias and variance. LMO consistently achieves smaller variance than LMR, which is consistent with the theoretical result of Section 3.2. Table 4 shows that, when the model assumption is incorrect, the attempt to capture nonlinear effects in linear terms result in biased estimates for the model-based approach. Here, the randomization-based approach is the better choice resulting in estimates much closer to the true value. Last, notice that $\text{INR}^{0.6}$ is more biased compared to SSR and in general is worse as a point estimator. How-

[8]To avoid problems with network density we test on $\text{INR}^{0.0} \equiv$ SSR, and $\text{INR}^{0.6}$. For example, running $\text{INR}^{1.0}$, typically, sets too many nodes in control for the networks we tested, and thus, has an adverse effect on manipulability. Optimal selection of the INR level is part of ongoing research.

Table 3. Results on the linear response function, listed as mean(σ)

Graphs			Estimation strategy			
Type	Parameters	Truth: δ_4	SSR	$INR^{0.6}$	LMR	LMO
Small world	$p = 0.05$	11.18	10.49(2.04)	9.15(2.36)	10.85(1.09)	9.7(0.63)
	$p = 0.5$	11.45	11.45(1.82)	10.60(2.60)	11.31(1.08)	11.66(0.76)
	$p = 0.9$	12.97	12.54(1.35)	12.17(1.57)	12.82(1.54)	12.42(0.72)
4-community block model	diag(0.9)/off(0.1)	3.19	3.47(0.59)	3.34(0.61)	3.15(0.56)	3.37(0.26)
	diag(0.25)/off(0.75)	3.28	3.56(0.74)	3.77(0.82)	3.28(0.27)	3.72(0.23)
	Beta(0.1,0.1)	3.19	3.2(0.49)	3.51(0.57)	3.25(0.36)	3.28(0.21)
	Beta(1,1)	3.27	3.41(0.57)	3.6(0.65)	3.30(0.33)	3.40(0.22)
Chung-Lu	-	2.88	2.9(0.61)	2.96(0.5)	2.99(0.31)	2.93(0.21)

Table 4. Results on the quadratic response function, listed as mean(σ)

Graphs			Estimation strategy			
Type	Parameters	Truth: δ_4	SSR	$INR^{0.6}$	LMR	LMO
Small world	$p = 0.05$	42.46	36.38(10.7)	32.13(13.96)	25.37(3.96)	27.32(1.49)
	$p = 0.5$	44.68	45.73(10.01)	39.52(14.24)	27.27(4.15)	35.45(2.05)
	$p = 0.9$	50.52	47.54(7.00)	46.87(9.12)	33.39(4.38)	36.00(1.80)
4-community block model	diag(0.9)/off(0.1)	5.48	5.97(1.25)	5.77(1.45)	6.64(0.9)	9.28(0.66)
	diag(0.25)/off(0.75)	5.63	6.01(1.26)	7.05(1.81)	8.96(1.20)	11.25(0.61)
	Beta(0.1,0.1)	5.42	5.63(0.94)	6.0(1.37)	8.40(1.19)	9.24(0.52)
	Beta(1,1)	5.59	6.07(1.08)	6.52(1.43)	6.53(0.75)	9.01(0.46)
Chung-Lu	-	4.65	4.83(0.75)	4.9(0.87)	6.87(0.74)	6.69(0.34)

ever, it generally achieves higher manipulability (see Section 4.1).

5. Concluding remarks

Adopting the potential outcomes framework for causal inference, we define a novel k-level estimand for peer influence effects and propose a randomization-based and a model-based approach to estimate it. Our randomization, namely INR, is a simple generalization of a sequential randomized design. INR aims to get more causal information (increase manipulability; see Table 2) at the expense of increased bias (see Tables 3,4), especially in dense networks.

The model-based approach performs efficient causal estimation in the presence of network uncertainty, when the additivity assumption holds. Furthermore, the model informs optimal assignment through maximizing Fisher information.

This work is a preliminary version of our research in the causal estimation of network effects, and it focuses mainly on introducing the problem and highlighting its conceptual challenges. A more refined version is forthcoming (Airoldi et al., 2013). Our future extensions include a formal statistical analysis of estimators arising from sequential randomizations (such as INR) and a more nuanced Bayesian analysis for optimal experimental design. Finally, we are actively applying our theory in two concrete problems that involve (i) peer influence in medical treatment compliance and (ii) interactions between modules in distributed computing systems.

Acknowledgments

The authors wish to express gratitude to Edoardo M. Airoldi and Donald B. Rubin for extensive discussions and support. This work was also supported by the Google U.S./Canada Fellowship in Statistics (2012-2013) and the Department of Defense (DoD) via the National Defense Science & Engineering Graduate Fellowship (NDSEG).

References

Acemoglu, Daron, Ozdaglar, Asuman, and ParandehGheibi, Ali. Spread of (mis)information in social networks. *Games and Economic Behavior*, 70(2): 194–227, 2010.

Airoldi, Edo, Toulis, Panos, Kao, Edward, and Rubin, Donald B. Estimation of causal peer influence effects. *Forthcoming*, 2013.

Aronow, P. M. and Samii, C. Estimating average causal effects under general interference. *working paper*, 2013.

Bakshy, Eytan, Eckles, Dean, Yan, Rong, and Rosenn, Itamar. Social influence in social advertising: Evidence from field experiments. *CoRR*, abs/1206.4327, 2012.

Bond, Robert M, Fariss, Christopher J, Jones, Jason J, Kramer, Adam DI, Marlow, Cameron, Settle, Jaime E, and Fowler, James H. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298, 2012.

Butts, C.T. Inference, error, and informant (in)accuracy. *Social Networks*, 25:103–140, 2003.

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

Manski, C.F. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993.

Mednick, S.C., Christakis, N.A., and Fowler, J.H. The spread of sleep behavior influences drug use in adolescent social networks. *PLoS One*, 5(3):e9775, 2010.

Ostrovsky, M. and Schwarz, M. Reserve prices in internet advertising auctions: A field experiment. *Available at SSRN 1573947*, (2054), 2010.

Parker, B.M. Design of network experiments. available at http://www.newton.ac.uk/programmes/DAE/seminars/090111301.pdf, 2011.

Pearl, Judea. *Causality: models, reasoning and inference*, volume 29. Cambridge Univ Press, 2000.

Perry, P.O. and Wolfe, P.J. Point process modeling for directed interaction networks. In *arXiv:1011.1703v1 [stat.ME]*, 2010.

Rosenbaum, P.R. Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200, 2007.

Rubin, D.B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology; Journal of Educational Psychology*, 66(5):688, 1974.

Rubin, D.B. [on the application of probability theory to agricultural experiments. essay on principles. section 9.] comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, 5(4):472–480, 1990.

Shah, Devavrat and Zaman, Tauhid. Rumors in a network: Who's the culprit? *IEEE Transaction on Information Theory*, 57(8):5163–5181, 2011.

Shalizi, C. and Thomas, A. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods and Research*, 40:211–239, 2011.

Sobel, M.E. What do randomized studies of housing mobility demonstrate? *Journal of the American Statistical Association*, 101(476):1398–1407, 2006.

Spirtes, Peter, Glymour, Clark, and Scheines, Richard. *Causation, prediction, and search*, volume 81. MIT press, 2001.

Tchetgen, E.J.T. and VanderWeele, T.J. On causal inference in the presence of interference. *Statistical Methods in Medical Research*, 21(1):55–75, 2012.

Ugander, J., Karrer, B., Backstrom, L., and Kleinberg, J. Network exposure to multiple universes. In *NIPS Workshop: Social network and social media analysis: Methods, models and applications*, Lake Tahoe, NV, 2012.

William, Aiello, Chung, Fan, and Lu, Linyuan. A random graph model for power law graphs. *Experimental Mathematics*, 10(1):53–66, 2001.

Zachary, W.W. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, pp. 452–473, 1977.