Causal Analysis

Impact Evaluation and Causal Machine Learning with Applications in R

Chapter 11: Treatment Evaluation under Interference Effects

Table of Contents

Failure of the Stable Unit Treatment Value Assumption

- So far, interference and spillover effects were ruled out.
- Assumed that outcome of any subject is not affected by the treatment status of any other subject, as postulated by the Stable Unit Treatment Value Assumption (SUTVA).
- However, SUTVA may appear unrealistic in many empirical problems, see Heckman, Lochner, and Taber (1998).

Example

- Training program: Share of individuals receiving a training (e.g. IT course) in a region may affect others' employment beyond their own training status, due to an increased regional supply of skills (IT competencies).
- Educational intervention: Provision of free textbooks may generate spillovers if treated students share the books with peers who did not directly receive them.

Disentangle Direct and Indirect Effects

Notation:

- Treatment of subject i is $D_i = d_i$.
- Treatments of all other subjects is $\mathcal{D}_{-i} = d_{-i}$.

Interference:

- Potential outcome of subject i given own and others' treatments is $Y_i(d_i, d_{-i}) \Rightarrow$ depends on others' treatments.
- Notation appears similar to dynamic treatments or mediation, but own and others' treatments need not be sequential.
- Goal is to separate the individual (or direct) effect of $D_i = d_i$ from the interference effect of $\mathcal{D}_{-i} = d_{-i}$.
- Interference complicates causal analysis, in particular if arbitrary forms of interference are allowed, see Manski (2013).
- Evaluations assessing interference and direct effects typically rely on assumptions about how interference affects the outcome.

Partial Interference and Exposure Mappings

Partial interference:

- Assume that interference effects are constrained to occur only within but not across specific (and nonoverlapping) clusters.
- In other words, SUTVA might be violated on an individual level, but holds on a cluster level.

Exposure mappings:

- Impose assumptions on the mechanisms through which interference affects outcomes based on information about the network of peers with which an individual interacts.
- For instance, assume that interference arises only from an individual's family members, while the treatment status of other individuals is irrelevant.

Table of Contents

Partial Interference and Double Randomization

Partial interference (Sobel, 2006; Hong and Raudenbush, 2006; and Hudgens and Halloran, 2008):

- SUTVA holds across clusters $\Rightarrow$ subjects' potential outcomes depend only on other subjects' treatments within clusters.

Identification through double randomization:

- Potential outcome of individual i in cluster c: $Y_{c,i}(d_{c,i}, d_{c,-i})$.
- Assumptions:
  - 1. Randomized treatment intensity on the cluster level, i.e., the treatment share in a specific cluster.
  - 2. Randomized treatment on the individual level, i.e., treatment assignment within cluster with a particular treatment share.
- This double randomization allows identifying the direct, interference, and total effects of a treatment.

Formal Discussion of Double Randomization

- Treated proportion within cluster c: $P_c = E[D|cluster = c]$.
- Under randomization of the proportion across clusters, any cluster's treated proportion and subjects' potential outcomes in that cluster are independent:

$$Y_{c,i}(d_{c,i}, \mathbf{d}_{c,-i}) \perp P_c \text{ for } d_{c,i} \in \{0,1\},$$
$$\text{any assignment } \mathbf{d}_{c,-i}, \text{ and any unit } i \text{ in any cluster } c. \tag{11.1}$$

- Under randomization of treated units within a cluster, individuals' potential outcomes are independent of individuals' treatment assignments within that cluster:

$$Y_{c,i}(d_{c,i}, \mathbf{d}_{c,-i}) \perp (D_{c,i}, \mathcal{D}_{c,-i}) | P_c \text{ for } d_{c,i} \in \{0,1\},$$
$$\text{any assignment } \mathbf{d}_{c,-i} \text{ and any unit } i \text{ in any cluster } c. \tag{11.2}$$

Direct and Interference Effect

- Difference in mean potential outcomes when varying individual treatment assignment while fixing the treatment share corresponds to the direct effect.
- Average direct effect of the individual treatment D_i for a given treatment share $P_c = p$:

$$\theta(p) = E[Y_{c,i}(1, p) - Y_{c,i}(0, p)]. \tag{11.4}$$

- Impact of shifting the treatment proportion from p to p' while fixing the individual treatment assignment corresponds to the interference effect.
- Average interference effect of p versus p' for a given $D_{c,i} = d$:

$$\delta(d, p, p') = E[Y_{c,i}(d, p) - Y_{c,i}(d, p')]. \tag{11.5}$$

Total Treatment Effect

- Total average treatment effect is the sum of the direct and interference effects:

$$\begin{aligned}\Delta(p,p') &= E[Y_{c,i}(1,p) - Y_{c,i}(0,p')] \\ &= \underbrace{E[Y_{c,i}(1,p) - Y_{c,i}(0,p)]}_{\theta(p)} + \underbrace{E[Y_{c,i}(0,p) - Y_{c,i}(0,p')]}_{\delta(0,p,p')} \\ &= \underbrace{E[Y_{c,i}(1,p') - Y_{c,i}(0,p')]}_{\theta(p')} + \underbrace{E[Y_{c,i}(1,p) - Y_{c,i}(1,p')]}_{\delta(1,p,p')} \qquad (11.5)\end{aligned}$$

- $\Delta(p,p')$ is the total effect among individuals switching the treatment status from 0 to 1 when increasing the treatment proportion from p' to p (these individuals are representative of the total population by double randomization).
- This total effect can be disentangled into direct and interference effects as shown in lines 2 and 3.

Overall Treatment Effect

- Overall treatment effect as suggested by Hudgens and Halloran (2008) is yet another causal parameter.
- Overall treatment effect is the average aggregate effect of assigning treatment proportions p versus p' to a population.
- Consider the potential (individual) treatment under treatment proportion p: $D_{c,i}(p)$.
- Overall treatment effect:

$$\tilde{\Delta}(p, p') = E[Y_{c,i}(D_{c,i}(p), p) - Y_{c,i}(D_{c,i}(p'), p')]. \tag{11.6}$$

- Being defined based on potential individual treatment states $D_{c,i}(p)$, $\tilde{\Delta}(p, p')$ acknowledges that some individual treatments might not change when shifting the treatment proportion.
- There may be always or never takers of the treatment in the notation of chapter 6.

Identification

Conditional mean differences of the direct, interference, total and overall effect are identified under double randomization:

$$\begin{aligned}
\theta(p) &= E[Y_{c,i}|D_{c,i}=1, P_c=p] - E[Y_{c,i}|D_{c,i}=0, P_c=p], \\
\delta(d,p,p') &= E[Y_{c,i}|D_{c,i}=d, P_c=p] - E[Y_{c,i}|D_{c,i}=d, P_c=p'], \\
\Delta(p,p') &= E[Y_{c,i}|D_{c,i}=1, P_c=p] - E[Y_{c,i}|D_{c,i}=0, P_c=p'], \\
\tilde{\Delta}(p,p') &= E[Y_{c,i}|P_c=p] - E[Y_{c,i}|P_c=p']. \qquad (11.7)
\end{aligned}$$

- Similar to mediation analysis, there might be interaction effects between direct and interference effects.
- Interference effects might differ across treatment states and the direct effect might depend on the treatment proportion.

Statistical Inference

Statistical inference methods:

- Hudgens and Halloran (2008) suggest variance estimators for the estimates of the causal effects in equation (11.7).
- Their approach assumes that interference effects depend only on the treatment proportion, not on who receives the treatment within a cluster (stratified interference assumption).
- Tchetgen and VanderWeele (2012) and Liu and Hudgens (2014) provide alternative methods for statistical inference without relying on the stratified interference assumption.

Interference Effects Example (1)

Example

Consider the impact of a training program for job seekers (D) on employment (Y):

- In some clusters, 50% of job seekers participate while in others, 0% (nobody) participate.
- The interference effect under individual nontreatment is the average effect of 50% in a cluster receiving training while individual i does not:

$$\delta(0, 0.5, 0) = E[Y_{c,i}|D_{c,i} = 0, P_c = 0.5] - E[Y_{c,i}|D_{c,i} = 0, P_c = 0] \quad (11.8)$$

- The average direct effect of individual training when 50% in a cluster receive training is given by:

$$\theta(0.5) = E[Y_{c,i}|D_{c,i} = 1, P_c = 0.5] - E[Y_{c,i}|D_{c,i} = 0, P_c = 0.5] \quad (11.9)$$

Interference Effects Example (2)

Example

- The total effect of being trained and exposed to an increase in the treatment proportion from 0 to 0.5 is given by:

$$\Delta(0.5, 0) = E[Y_{c,i}|D_{c,i} = 1, P_c = 0.5] - E[Y_{c,i}|D_{c,i} = 0, P_c = 0] \quad (11.10)$$

- The overall treatment effect is:

$$\tilde{\Delta}(0.5, 0) = E[Y_{c,i}|P_c = 0.5] - E[Y_{c,i}|P_c = 0] \quad (11.11)$$

- However, assessing the interference effect under individual treatment $\delta(1, 0.5, 0)$ and the direct effect $\theta(0)$ is impossible.
- The reason is that the lower treatment proportion of zero implies that there are no treated individuals under $P_c = 0$.

IV Approach to Partial Interference

Alternatives to double randomization:

- Besides double randomization, there exist alternative strategies such as IV, DiD, and RDD approaches.

Instrumental variable approach:

- If some individuals do not comply with their random treatment assignment, an IV approach might be considered.
- Random individual treatment assignment instruments actual treatment participation.
- Requires interference-adjusted IV assumptions, as discussed by Kang and Imbens (2016) and Imai, Jiang, and Malani (2021).
- Direct, interference, and total effects can be estimated dividing the intention-to-treat (ITT) effect by the first-stage effect.

Selection-on-Observables Approach to Partial Interference

Selection-on-observables approach:

- Assume that equations (11.1) and (11.2) only hold conditional on cluster-specific covariates X_c.
- Assume further that only treatment proportion drives the interference effects, as in Ferracci, Jolivet, and van den Berg (2014).
- Direct, interference, total, and overall effects given X_c are identified by adding X_c as conditioning set in expression (11.7), see VanderWeele (2010).
- For instance, average direct effect conditional on X_c is:

$$\begin{aligned}\theta_{X_c}(p) &= E[Y_{c,i}(1,p) - Y_{c,i}(0,p)|X_c] \\ &= E[Y_{c,i}|D_{c,i}=1, P_c=p, X_c] - E[Y_{c,i}|D_{c,i}=0, P_c=p, X_c] \quad (11.12)\end{aligned}$$

Cluster Randomization Approach to Partial Interference (1)

Random treatment availability across clusters:

- Randomly assign treatment availability across clusters, while individual treatment depends on explicit, nonrandom eligibility criterion, see e.g. Angelucci and De Giorgi (2009).
- For example, welfare program for poor households is randomly assigned across geographic regions.
- In treated regions, only households below a certain poverty level are eligible for the welfare program.
- In nontreated regions, the welfare program is not available for any household.

Cluster Randomization Approach to Partial Interference (2)

- Assume eligibility is a known, deterministic function of (an) observed variable(s) and defined as binary indicator $\mathcal{E}$.
- P_c is a binary variable indicating whether treatment is available in a cluster, such as a region.
- Define the interference and total treatment effect conditional on the eligibility status:

$$\begin{aligned} \delta_{\mathcal{E}=e}(d) &= E[Y_{c,i}(d,1) - Y_{c,i}(d,0) \mid \mathcal{E}_{c,i} = e], \\ \Delta_{\mathcal{E}=e} &= E[Y_{c,i}(1,1) - Y_{c,i}(0,0) \mid \mathcal{E}_{c,i} = e]. \end{aligned} \tag{11.13}$$

- Interference effect on ineligible subjects is identified:

$$\delta_{\mathcal{E}=0}(0) = E[Y_{c,i} \mid \mathcal{E}_{c,i} = 0, P_c = 1] - E[Y_{c,i} \mid \mathcal{E}_{c,i} = 0, P_c = 0]. \tag{11.14}$$

- Total effect on eligible subjects is identified:

$$\Delta_{\mathcal{E}=1} = E[Y_{c,i} \mid \mathcal{E}_{c,i} = 1, P_c = 1] - E[Y_{c,i} \mid \mathcal{E}_{c,i} = 1, P_c = 0]. \tag{11.15}$$

DiD Approach to Partial Interference (1)

Difference-in-differences approach:

- Requires both pretreatment and posttreatment periods.
- Maintain assumption that eligibility status $\mathcal{E}$ is observed, but replace assumption (11.1) with a common trend assumption.
- Conditional on eligibility status, the mean potential outcome among the treated clusters in absence of the treatment would follow the same time trend as the nontreated clusters.
- Considering time index T to indicate the pretreatment period ($t = 0$) and posttreatment period ($t = 1$) and the potential outcome $Y_{c,i,t}(d, P_c)$, the common trend assumption is:

$$\begin{aligned} &E[Y_{c,i,1}(0,0) - Y_{c,i,0}(0,0) | \mathcal{E}_{c,i} = e, P_c = 1] \\ &= E[Y_{c,i,1}(0,0) - Y_{c,i,0}(0,0) | \mathcal{E}_{c,i} = e, P_c = 0] \end{aligned} \tag{11.16}$$

DiD Approach to Partial Interference (2)

- As discussed in Huber and Steinmayr (2021), the common trend assumption in equation (11.16) and the identifiably of eligibility types permits identifying the following effects.
- $\Delta_{\mathcal{E}=1,P_c=1}$: total effect on eligible subjects in treated clusters.
- $\delta_{\mathcal{E}=0,P_c=1}(0)$: interference effect under nontreatment on ineligible subjects in treated clusters.

$$
\begin{aligned}
\Delta_{\mathcal{E}=1,P_c=1} &= E[Y_{c,i,1}|\mathcal{E}_{c,i}=1,P_c=1] - E[Y_{c,i,0}|\mathcal{E}_{c,i}=1,P_c=1] \\
&\quad - \{E[Y_{c,i,1}|\mathcal{E}_{c,i}=1,P_c=0] - E[Y_{c,i,0}|\mathcal{E}_{c,i}=1,P_c=0]\}, \\
\delta_{\mathcal{E}=0,P_c=1}(0) &= E[Y_{c,i,1}|\mathcal{E}_{c,i}=0,P_c=1] - E[Y_{c,i,0}|\mathcal{E}_{c,i}=0,P_c=1] \\
&\quad - \{E[Y_{c,i,1}|\mathcal{E}_{c,i}=0,P_c=0] - E[Y_{c,i,0}|\mathcal{E}_{c,i}=0,P_c=0]\} \qquad (11.17)
\end{aligned}
$$

RDD Approach to Partial Interference

Regression discontinuity design (RDD) approach:

- RDD exploits a running variable at the cluster level with a threshold determining treatment availability in a cluster P_c, see Angelucci and Di Maro (2016).
- Example: Regional development fund for financing local businesses, which is provided depending on whether the regional GDP exceeds a certain threshold.
- Regions just above and below the threshold are presumably comparable.
- If businesses have to meet eligibility criteria to receive funding, equations (11.14) and (11.15) identify effects locally around the threshold.

Further Approaches to Partial Interference

Further potential strategies:

- IV methods using a cluster-level instrument for P_c.
- Selection-on-observables approach by assuming that expression (11.1) holds conditional on covariates X_c.
- Selection on observables implies that equations (11.15) and (11.14) given X_c yield the respective causal parameters conditional on covariates.
- In contrast, Forastiere, Mealli, and VanderWeele (2016) assume only a subset of potential outcomes (e.g., under individual treatment) are independent of eligibility conditional on X_c.
- Full identifiability of eligibility not required, but effect identification for specific eligibility/compliance types (based on individual treatment reaction to regional treatment availability) requires additional assumptions, e.g. effect homogeneity.

Table of Contents

Introduction to Exposure Mappings

Interference based on exposure mappings:

- In contrast to partial interference (which considers interference within clusters but is agnostic about networks within clusters), exposure mappings focus on and impose restrictions on a subject's relevant interference network.
- For this reason, exposure mappings require information on subjects' networks, namely contacts they interact with, which determines the strength of interference.
- For example, an information campaign may not only directly affect someone's opinion, but also the opinions of her or his contacts in the network, which in turn may also exert an interference effect on a subject's own opinion through social interactions.

Restricting Complexity

- Exposure mappings aim to restrict the (otherwise excessive) complexity of possible interference effects.

Possible restrictions:

- Interference effects depend only on subjects' number of contacts but not on quality (family or friends).
- Interference effects depend only on whether at least one contact is treated (exact number of treated contacts is irrelevant).
- Based on such restrictions, exposure mappings allow defining multiple (but not an excessive number of) kinds of interference.

Controlling for Confounding

- Evaluate direct and interference effects by controlling for differences in probabilities of exposure mappings across subjects, particularly due to different networks.
- Network features like the number of contacts are generally confounders (affect exposure mappings and outcome), even if individual treatment is randomized.
- For instance, subjects with larger networks are more likely to have at least one treated contact and may also differ in terms of their outcomes (like opinions) from subjects with smaller networks.

Direct, Indirect, and Total Effects

- Exposure mapping $\mathcal{G}_i$ for individual i is denoted by:

$$\mathcal{G}_i = \mathcal{F}(\mathcal{N}_i, \mathcal{D}_{-i}), \tag{11.18}$$

  where $\mathcal{N}_i$ is individual i's interference network,
  $\mathcal{D}_{-i}$ is other subjects' treatment assignment, and
  $\mathcal{F}$ is a presumably known function.

- When correctly assuming the exposure mapping, $Y_i(d_i, \mathbf{d}_{-i})$ simplifies to $Y_i(d, g)$ such that direct, interference, and total effects under two exposure mappings g and g' become:

$$\begin{aligned} \theta(g) &= E[Y_i(1, g) - Y_i(0, g)], \\ \delta(d, g, g') &= E[Y_i(d, g) - Y_i(d, g')], \\ \Delta(g, g') &= E[Y_i(1, g) - Y_i(0, g')] \end{aligned} \tag{11.19}$$

Random Treatment Assignment

- Assume random treatment assignment among individuals in the population after the social networks have been formed:

$$Y_i(d_i, \mathbf{d}_{-i}) \perp (D_i, \mathcal{D}_{-i}) \text{ for } d_i \in \{0, 1\} \text{ and any assignment } \mathbf{d}_{-i} \quad (11.20)$$

- Despite random treatment, exposure mappings are not random.
- Need to control for joint probability of i's potential treatment state D_i and exposure mapping $\mathcal{G}_i$ as a function of network structure $\mathcal{N}_i$, denoted by $p_i(d, g) = \Pr(D_i = d, \mathcal{G}_i = g | \mathcal{N}_i)$.
- Inverse probability weighting yields direct and interference effects, see Aronow and Samii (2017), who also provide conservative inference method:

$$\theta(g) = E\left[\frac{Y_i \cdot D_i \cdot I\{\mathcal{G}_i = g\}}{p_i(1, g)} - \frac{Y_i \cdot (1 - D_i) \cdot I\{\mathcal{G}_i = g\}}{p_i(0, g)}\right],$$

$$\delta(d, g, g') = E\left[\frac{Y_i \cdot I\{D_i = d\} \cdot I\{\mathcal{G}_i = g\}}{p_i(d, g)} - \frac{Y_i \cdot I\{D_i = d\} \cdot I\{\mathcal{G}_i = g'\}}{p_i(d, g')}\right] \quad (11.21)$$

Selection on Observables

If treatment assignment of i's treatment D_i and exposure mapping $\mathcal{G}_i$ is as good as random given covariates X and network structure $\mathcal{N}_i$:

- Inverse probability weighting controlling for individual- and network-related covariates when computing the propensity score $p_i(d, g)$.
- Matching or regression to adjust propensity scores, as suggested by Forastiere, Airoldi, and Mealli (2021).
- Doubly robust estimation like targeted maximum likelihood estimation as discussed by Van der Laan (2014).
- For doubly robust estimation that combines exposure mappings and partial interference, see Qu, Xiong, Liu, and Imbens (2021).
- See Aronow, Eckles, Samii, and Zonszein (2020) for a survey on causal analysis under interference.