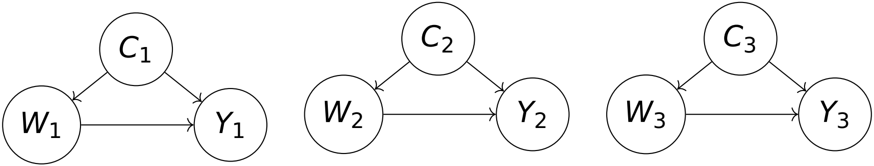
Treatment Effect Estimation with Interference and Confounding using Graphical Models

Leonard Henckel, Meta-Lina Spohn and Marloes H. Maathuis

Interference

Stable unit treatment value assumption (SUTVA):
(Potential) outcome on one unit should be unaffected by treatment assignments of other units



- No interference: W_i does not affect Y_j for $i \neq j$
- No contagion: Y_i does not affect Y_j for $i \neq j$

Interference

Examples with interference:

- Crosscontamination in agricultural studies
- Instruction type in school
- Vaccinations
- Communication of advertisement

Remark: Common were units can communicate/compete

Interference

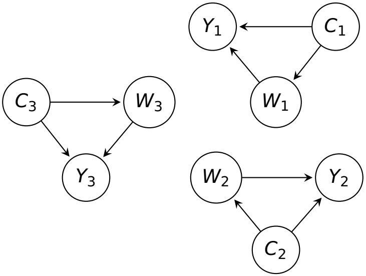
No interference

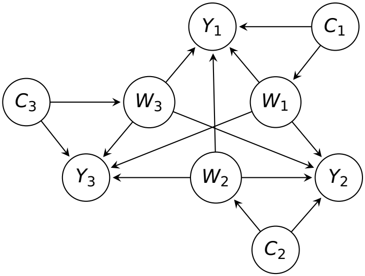
- $C_i \leftarrow \epsilon_{C_i}$
- $W_i \leftarrow g(C_i, \epsilon_{W_i})$
- $Y_i \leftarrow h(W_i, C_i, \epsilon_i)$

Interference

- $C_i \leftarrow \epsilon_{C_i}$
- $W_i \leftarrow g(C_i, \epsilon_{W_i})$
- $Y_i \leftarrow h(\boldsymbol{W}, C_i, \epsilon_i)$,

$\boldsymbol{W} := (W_1, \dots, W_N)$.





Causal effects

Classical i.i.d. setting:

- Two possible treatment assignments
- Average Treatment Effect:
  $E[Y|do(W=1)] - E[Y|do(W=0)]$

Interference setting:

- 2^N possible treatment assignments
- Global Average Treatment Effect:
  $\tau = \frac{1}{N}\sum_{i=1}^{N}\left(E[Y_i|do(\boldsymbol{W}=\mathbf{1})] - E[Y_i|do(\boldsymbol{W}=\mathbf{0})]\right)$
- Stochastic Interventions:
  $\tau(P_\pi, P_\eta) = \frac{1}{N}\sum_{i=1}^{N}\left(E[Y_i|do(\boldsymbol{W}\sim P_\pi)] - E[Y_i|do(\boldsymbol{W}\sim P_\eta)]\right)$

Literature on interference

Conceptual settings:

- Interference as nuisance to correct for
- Interference as part of effect

Approaches to handling interference:

- Experimental design (e.g. VanderWeele and Tchetgen Tchetgen, 2011)
- Partial interference: Arbitrary interference but within non-overlapping groups (e.g. Hudgens and Halloran, 2008)
- Explicit modelling of interference (e.g. Chin, 2019)

Remark: Without restrictions on the interference, treatment effects are generally not identifiable.

Modelling interference with features

Network graph G:
Graph describing interactions between units, e.g,

- Friendship ties in a social group
- Geographical adjacency

Features: $\boldsymbol{X} = f(\mathbf{W}, G)$ (Manski, 1993; Chin, 2019)
Deterministic functions of G and treatment vector $\mathbf{W}$, e.g,

- Fraction of treated neighbors
- Fraction of treated neighbors of neighbors

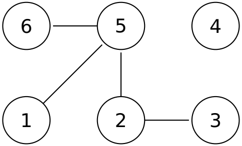
Modelling interference:

$$Y_i \leftarrow h(\boldsymbol{W}, \boldsymbol{C}_i, \epsilon_i) = h'(W_i, \boldsymbol{X}_i, \boldsymbol{C}_i, \epsilon_i)$$

Modelling interference with features

Example: exposure to advertisement

- $X_i = f(\boldsymbol{W}, G) =$ number of treated neighbors of i in G,
- $Y_i \leftarrow \mu_i + W_i + X_i + \epsilon_i$.



$$\tau = \frac{1}{6}\sum_{i=1}^{6}\left(E[Y_i|do(\boldsymbol{W}=\mathbf{1})] - E[Y_i|do(\boldsymbol{W}=\mathbf{0})]\right)$$

$$= \frac{1}{6}\sum_{i=1}^{6}\left(1 + E[X_i|do(\boldsymbol{W}=\mathbf{1})] - 0 - E[X_i|do(\boldsymbol{W}=\mathbf{0})]\right),$$

where $E[X_i|do(\boldsymbol{W}=\mathbf{1})] = (1, 2, 1, 0, 3, 1)$.

$\implies \tau = (2+3+2+1+4+2)/6 - 0 = 14/6$.

Model

Definition

Unit SEM with interference features & linear outcome model (inspired by Chin, 2019):

- $\boldsymbol{C}_i \leftarrow \mathcal{A}\boldsymbol{C}_i + \epsilon_{C_i}$,
- $W_i \leftarrow g(\boldsymbol{C}, \epsilon_{W_i})$,
- $\boldsymbol{X}_i \leftarrow f(\boldsymbol{W}, G)$,
- $Y_i \leftarrow W_i\beta^{(1)}(1, \boldsymbol{X}_i) + (1 - W_i)\beta^{(0)}(1, \boldsymbol{X}_i) + \gamma\boldsymbol{C}_i + \epsilon_i$.

Remark: Outcome model corresponds to

$$Y_i \leftarrow \alpha^{(0)}(1, \boldsymbol{X}_i) + \alpha^{(1)}(W_i, \boldsymbol{O}_i) + \gamma\boldsymbol{C}_i + \epsilon_i,$$

with $\beta^{(0)} = \alpha^{(0)}, \beta^{(1)} = \alpha^{(0)} + \alpha^{(1)}$ and $\boldsymbol{O}_i = W_i\boldsymbol{X}_i$.

Results overview

Results:

- Reformulate GATE τ and stochastic effects $\tau(P_\pi, P_\eta)$
- Define new notion of Interference-DAG (I-DAG)
- Propose adjustment estimator based on graphical criteria on I-DAG
- Propose instrumental variable estimator based on graphical criteria on I-DAG
- Derive sufficient conditions for consistency
- Derive sufficient conditions for asymptotic normality

GATE reformulation

Lemma: (Stochastic treatment contrast with features)

$\tau_N(P_\pi, P_\eta) = \boldsymbol{\omega}_N^{(1)}(\pi,\eta)(\boldsymbol{\alpha}^{(0)} + \boldsymbol{\alpha}^{(1)}) - \boldsymbol{\omega}_N^{(0)}(\pi,\eta)\boldsymbol{\alpha}^{(0)}$ where

$$\boldsymbol{\omega}_N^{(1)} = \frac{1}{N}\sum_{i=1}^{N}\left(\pi E[X_i | do(W \sim P_\pi)] - \eta E[X_i | do(W \sim P_\eta)]\right), \text{ and}$$

$$\boldsymbol{\omega}_N^{(0)} = \frac{1}{N}\sum_{i=1}^{N}\left((1-\pi) E[X_i | do(W \sim P_\pi)] - (1-\eta) E[X_i | do(W \sim P_\eta)]\right.$$

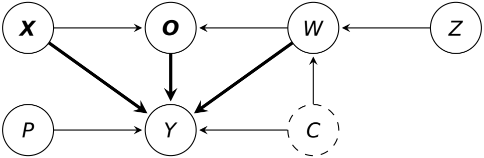
Remark: Weights can be computed or estimated with Monte Carlo so only need to estimate $\boldsymbol{\alpha}^{(0)}$ and $\boldsymbol{\alpha}^{(1)}$.

I-DAG

Idea: interpret feature model as non-i.i.d. data from DAG

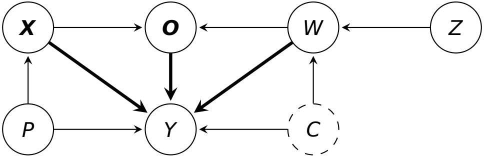
Example:

- $Z_i \leftarrow \epsilon_{Z_i}$, $P_i \leftarrow \epsilon_{P_i}$, $C_i \leftarrow \epsilon_{C_i}$,
- $W_i \leftarrow g(Z_i, \epsilon_{W_i})$,
- $\boldsymbol{X}_i \leftarrow f(\boldsymbol{W}_{-i}, G)$,
- $\boldsymbol{O}_i = W_i \boldsymbol{X}_i$,
- $Y_i = \alpha_0^{(0)} + \alpha_{1:P}^{(0)} \boldsymbol{X}_i + \alpha_0^{(1)} W_i + \alpha_{1:P}^{(1)} \boldsymbol{O}_i + \gamma_C \boldsymbol{C}_i + \gamma_P P_i + \epsilon_i$.



Example: Unit SEM:

- $Z_i \leftarrow \epsilon_{Z_i}$, $P_i \leftarrow \epsilon_{P_i}$, $C_i \leftarrow \epsilon_{C_i}$,
- $W_i \leftarrow g(Z_i, \boldsymbol{P}_{-i}, \epsilon_{W_i})$,
- $\boldsymbol{X}_i \leftarrow f(\boldsymbol{W}_{-i}, G)$,
- $\boldsymbol{O}_i = W_i \boldsymbol{X}_i$,
- $Y_i = \alpha_0^{(0)} + \alpha_{1:P}^{(0)} \boldsymbol{X}_i + \alpha_0^{(1)} W_i + \alpha_{1:P}^{(1)} \boldsymbol{O}_i + \gamma_C C_i + \gamma_P P_i + \epsilon_i$.



I-DAG theory

Definition (Interference-DAG (I-DAG))

- If $A_i \to B_i$ for some i in the unit graph, then $A \to B$ in the I-DAG.
- If $A_i \to W_j$ for some $i \neq j$ in the unit graph, then $A \to X$ in the I-DAG.

Lemma (I-DAG encodes independence statements)

If $\mathbf{A} \perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{C}$ *in the I-DAG, then* $\mathbf{A}_i \perp\!\!\!\perp \mathbf{B}_i \mid \mathbf{C}_i$ *for all* $i = 1, \ldots, N$.

I-DAG

Idea:

- $\boldsymbol{\alpha}^{(0)}$ and $\boldsymbol{\alpha}^{(1)}$ suffice to estimate $\tau_N(P_\pi, P_\eta) = \boldsymbol{\omega}_N^{(1)}(\pi, \eta)(\boldsymbol{\alpha}^{(0)} + \boldsymbol{\alpha}^{(1)}) - \boldsymbol{\omega}_N^{(0)}(\pi, \eta)\boldsymbol{\alpha}^{(0)}$.
- $(\boldsymbol{\alpha}^{(0)}, \boldsymbol{\alpha}^{(1)})$ is total effect of pseudo-intervention on $(W, \boldsymbol{X}, \boldsymbol{O})$ on Y in I-DAG $\mathcal{G}$.

  $\Longrightarrow$ Apply graphical criteria from literature to I-DAG to estimate $\boldsymbol{\alpha}^{(0)}$ and $\boldsymbol{\alpha}^{(1)}$

Lemma

$\mathbf{P} = \mathrm{pa}(Y, \mathcal{G}) \setminus \boldsymbol{T}$ *and* $\mathbf{Z}$ *valid adjustment set. Then* $\beta_{y\boldsymbol{t}.\mathbf{p}} = \beta_{y\boldsymbol{t}.\mathbf{z}}$ *where* $\mathbf{P} = \mathrm{pa}(Y, \mathcal{G}) \setminus \{X\}$ *and* $\boldsymbol{T} = \{W\} \cup \boldsymbol{X} \cup \boldsymbol{O}$*.*

Assumptions on feature behaviour

Assumption 1:
Assume existence of:

- $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[X_i | do(W \sim P_\pi)]$
- $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[X_i | do(W \sim P_\eta)]$.

Assumption 2:
Constrain dependencies between units by using yet another auxiliary graph called dependency graph.

Remark: Implicitly constrains underlying network graph without explicitly modelling it.

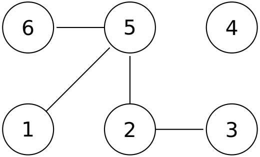
Dependency graph

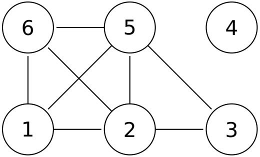
Definition: (Interference dependency graph)
$U_i \to U_j$ present in $D(\boldsymbol{X}, \boldsymbol{W})$ iff:

a) W_i affects $\boldsymbol{X}_j$,

b) $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ are affected by some W_k, $k \neq i, j$.

Example: X_i = number of treated neighbors of i in G.





Adjustment consistency

Theorem: (Consistency for adjustment estimator)
$\boldsymbol{D}$ valid adjustment set relative to $(\boldsymbol{A}, Y)$ with $\boldsymbol{A} = (\boldsymbol{X}, W, \boldsymbol{O})$ in I-DAG $\mathcal{G}$, $\tilde{\boldsymbol{X}} = (\boldsymbol{A}, \boldsymbol{D})^T$. If

- $E[Y_i^4] < \infty$ and $E\left[\tilde{\boldsymbol{X}}_i^4\right] < \infty \; \forall \; i = 1, \ldots, N$
- $E\left[\tilde{\boldsymbol{X}}_i \tilde{\boldsymbol{X}}_i^T\right] < \infty$ is invertible $\forall \; i = 1, \ldots, N$,
- $\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} E\left[\tilde{\boldsymbol{X}}_i \tilde{\boldsymbol{X}}_i^T\right] = \Sigma_{\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}} < \infty$, where $\Sigma_{\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}}$ is invertible,
- $d(N) \in o(N)$, where $d(N)$ is the maximal degree in the dependency graph $D(X, W)$,

then

$$\tau_{ols}(P_\pi, P_\eta) - \tau_N(P_\pi, P_\eta) \xrightarrow{P} 0,$$

Remark: $d(N) \in o(N) \implies$ no non-local features

Adjustment asymptotic normality

Theorem: (Asymptotic normality of adjustment estimator)
$\boldsymbol{D}$ valid adjustment set relative to $(\boldsymbol{A}, Y)$ with $\boldsymbol{A} = (\boldsymbol{X}, W, \boldsymbol{O})$ in I-DAG $\mathcal{G}$, $\tilde{\boldsymbol{X}} = (\boldsymbol{A}, \boldsymbol{D})^T$. If

- $E[Y_i^8] < \infty$ and $E\left[\tilde{\boldsymbol{X}}_i^8\right] < \infty \; \forall \; i = 1, \ldots, N$,
- $E\left[\tilde{\boldsymbol{X}}_i \tilde{\boldsymbol{X}}_i^T\right] < \infty$ is invertible $\forall \; i = 1, \ldots, N$,
- $\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N E\left[\tilde{\boldsymbol{X}}_i \tilde{\boldsymbol{X}}_i^T\right] = \Sigma_{\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}} < \infty$, where $\Sigma_{\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}}$ is invertible,
- $\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N E\left[\epsilon_i^2 \tilde{\boldsymbol{X}}_i \tilde{\boldsymbol{X}}_i^T\right] = \Sigma_{\epsilon^2 \tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}} < \infty$, where $\epsilon_i := Y_i - \tilde{\boldsymbol{X}}_i^T (\gamma_A, \gamma_D)^T$,
- $d(N) \in o(N^{1/4})$,

then

$$\sqrt{N} \left(\tau_{ols}(P_\pi, P_\eta) - \tau_N(P_\pi, P_\eta)\right) \xrightarrow{D} \mathcal{N}(0, \sigma^2),$$

Summary

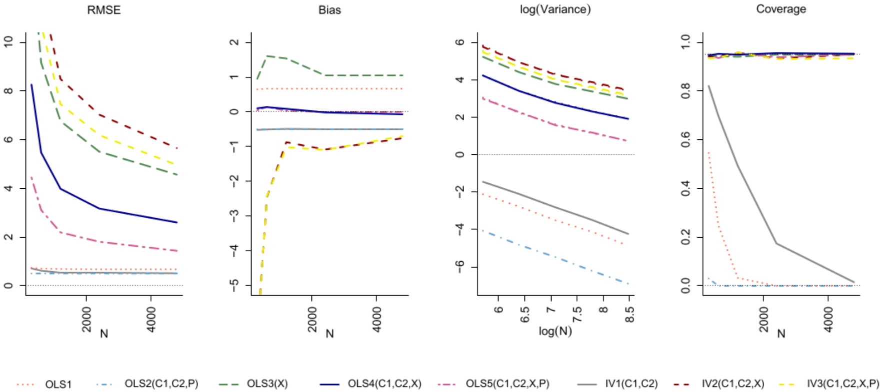
- Framework that allows for both confounding and interference
- I-DAG to make classical graphical results usable
- Adjustment estimator
- Instrumental variables estimator
- Necessary conditions for consistency and asymptotic normality

Thanks!

References:

- Spohn, M.-L., Henckel, L. and Maathuis, M. H. (2022) Treatment Effect Estimation on Networks with Interference and Confounding
- Pearl, J. (2009). Causality. *Cambridge University Press*.
- Chin, A. (2019). Regression adjustments for estimating the global treatment effect in experiments with interference. *Journal of Causal Inference*
- Henckel, L., Buttenschoen, M., and Maathuis, M. H. (2022). Graphical tools for efficient conditional instrumental set selection.
- Perković, E., Textor, J., Kalisch, M., and Maathuis, M. H. (2018). Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs. *Journal of Machine Learning Research*
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*
- Hudgens, M. and Halloran, M. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*
- VanderWeele, Tyler J., and Eric J. Tchetgen Tchetgen. Effect partitioning under interference in two-stage randomized vaccine trials. *Statistics & probability letters*

Simulations



$\mathcal{G}$: Watts-Strogatz network, Target: $\tau(P_\pi, P_\eta)$ with $\pi = 0.7$ and $\eta = 0.2$, $\boldsymbol{X}$: i) fraction of treated neighbors and ii) fraction of treated neighbors of neighbors.