Bayesian Causal Inference in the Presence of Endogenous Selection into Treatment and Spillovers

Duong Trinh

Research Summary

August 2024

Introduction

- ✗ ~~Stable Unit Treatment Value Assumption (SUTVA):~~

  ~~Units are affected by their *own* treatment status.~~
- ➥ Spillovers aka Interference:

  Units are also affected by the treatment status of *others*.

- ✗ ~~Unconfoundedness:~~

  ~~*No unobserved confounders* that affect both treatment assignment and the outcome.~~
- ➥ Endogenous selection into treatment:

  Treatment assignment is influenced by *unobserved factors* that also affect the outcome.

- ➥ I propose a model and Bayesian methods that use observational (network/spatial) data to estimate direct and indirect causal effects while taking account of unobserved heterogeneity.

Motivating Examples

1. **Q:** Causal effects of *Opportunity Zone (OZ) program* on *census tracts' economic development*?

   Endogenous selection: unobservable characteristics in OZ program assignment process may be related to regional economics performance.

   Spatial Interference: when OZ program is introduced in a census tract, surrounding areas not designated as OZs would also be affected.

2. **Q:** Causal effects of *SEL-focused after-school program* on *students' prosocial development*?

   Endogenous selection: participation is not mandatory; consequently, students who sign up may be inherently different from non-participants in ways that are related to the outcome being measured.

   Network Interference: enrolled-students might also interact with their peers who are not assigned to the program, leading to knowledge transfer or behavioral influence.

Related Literature I

1. Violation of SUTVA $\rightarrow$ Spillovers

Partial interference: spillovers occur within clusters but not across clusters.

Sobel (2006), Hudgens and Halloran (2008), Manski (2013), DiTraglia et al. (2023), etc.

$\rightarrow$ limited settings when units naturally cluster a significant distance apart.

General interference: using network/spatial data to define the form of spillovers.

Randomized experiments on social networks

Toulis and Kao (2013), Aronow and Samii (2017), Leung (2020), Yuan and Altenburger (2022), etc.

Observational studies

van der Laan and Sofrygin (2017), Forastiere et al. (2021), Forastiere et al. (2022), Ogburn et al. (2022), Sanchez-Becerra (2022), Leung and Loupos (2022), etc.

$\rightarrow$ reliance on unconfoundedness, little exploration of selection on unobservables.

Related Literature II

2. Endogenous Selection Models
   - Gaussian generalised Roy model (Li et al., 2004; Heckman et al., 2014), etc.
   - Sample selection models (Ding, 2014; Doğan and Taşpinar, 2018), etc.

Contributions

- ♣ New method to estimate direct and indirect causal effects in the presence of endogenous selection into treatment and spillovers.
- ♣ Key idea: explicitly models endogenous selection into treatment and employs network/spatial data to capture spillovers in the form of a neighbourhood treatment term.
  - allows for heterogeneous direct effects across individuals.
  - allows for general interference.
  - embraces Bayesian methods to facilitate straightforward estimation and inference and to relax parametric assumptions (*further steps*).

Outline

Outline

1. Model and Causal Estimands
2. Estimation and Inference
3. Simulation Study
4. Empirical Application
5. Conclusion

General Setup

- There are n agents:

$$i = 1, \ldots, n \tag{1}$$

- *Individual Treatment*
  $\sim$ Self-selection process:

$$D_i = \mathbb{1}\{\nu(Z_i, X_i) \geq \overbrace{V_i}^{\text{unobservable resistance to treatment}}\} \tag{2}$$

- *Neighborhood Treatment*
  $\sim$ A summary measure:

$$\bar{D}_{\mathcal{N}i} = \sum_{j=1, j\neq i}^{N} \overbrace{w_{ij}}^{\text{from adjacency matrix}} D_j \tag{3}$$

- *Treated* and *Untreated* Potential Outcomes:

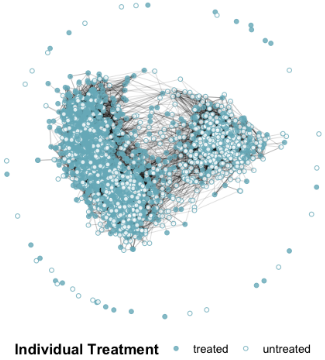
$$Y_i^{(1)} = \mu_1(\bar{D}_{\mathcal{N}i}, X_i) + \epsilon_i^{(1)} \tag{4}$$

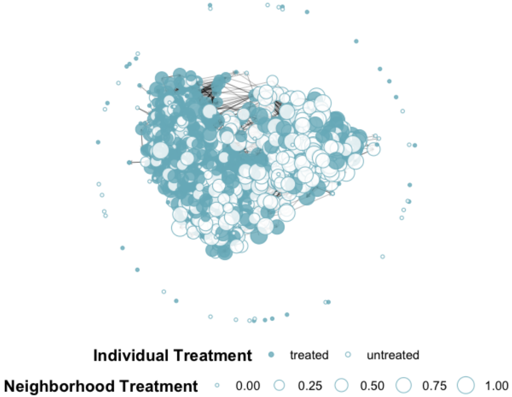
$$Y_i^{(0)} = \mu_0(\bar{D}_{\mathcal{N}i}, X_i) + \epsilon_i^{(0)} \tag{5}$$

- *Revealed* Outcome:

$$Y_i = D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)} \tag{6}$$

Exposure to the policy





The Model

♣ Parametric Model

$$D_i = \mathbb{1}\{\alpha Z_i + \beta^{(D)} X_i + \epsilon_i^{(D)} \geq 0\}$$

$$\bar{D}_{\mathcal{N}i} = \sum_{j=1, j\neq i}^{N} w_{ij} D_j$$

$$Y_i^{(1)} = \delta^{(1)} \bar{D}_{\mathcal{N}i} + \beta^{(1)} X_i + \epsilon_i^{(1)}$$

$$Y_i^{(0)} = \delta^{(0)} \bar{D}_{\mathcal{N}i} + \beta^{(0)} X_i + \epsilon_i^{(0)}$$

$$Y_i = D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}$$

♣ Assumptions

A1. $(\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)})^T \perp (X_i, Z_i)$

A2. $(\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)})^T \overset{ind}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & \sigma_{1D} & \sigma_{0D} \\ & \sigma_1^2 & \sigma_{10} \\ & & \sigma_0^2 \end{bmatrix}$$

A3. $Y_i^{(1)}(\bar{d}_{\mathcal{N}}) \perp \bar{D}_{\mathcal{N}i} \mid X_i \quad \forall \bar{d}_{\mathcal{N}} \in [0, 1]$
$Y_i^{(0)}(\bar{d}_{\mathcal{N}}) \perp \bar{D}_{\mathcal{N}i} \mid X_i \quad \forall \bar{d}_{\mathcal{N}} \in [0, 1]$

Causal Estimands

♣ Indirect Causal Effects

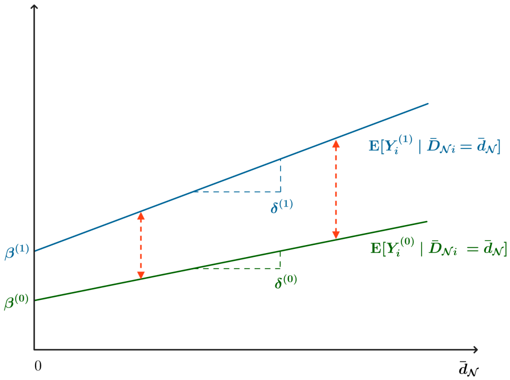
- $\bar{D}_{\mathcal{N}i} = \bar{d}_{\mathcal{N}} \to \bar{d}_{\mathcal{N}} + \Delta$; hold $D_i = d$ fixed.
- Average Indirect effect when untreated $AIE^{(1)}(\bar{d}_{\mathcal{N}}, \Delta) := \mathbb{E}\left[Y_i^{(1)}(\bar{d}_{\mathcal{N}} + \Delta) - Y_i^{(1)}(\bar{d}_{\mathcal{N}})\right]$
- Average Indirect effect when treated $AIE^{(0)}(\bar{d}_{\mathcal{N}}, \Delta) := \mathbb{E}\left[Y_i^{(0)}(\bar{d}_{\mathcal{N}} + \Delta) - Y_i^{(0)}(\bar{d}_{\mathcal{N}})\right]$

♣ Direct Causal Effects

- $D_i = 0 \to 1$; hold $\bar{D}_{\mathcal{N}i} = \bar{d}_{\mathcal{N}}$ fixed.
- Average Direct Treatment Effect

$$ADTE(\bar{d}_{\mathcal{N}}) := \mathbb{E}\left[Y_i^{(1)} - Y_i^{(0)} \mid \bar{d}_{\mathcal{N}}, X_i\right]$$

Causal Estimands



- Average Partial Indirect Effects

  Treated: $\delta^{(1)}$

  Untreated: $\delta^{(0)}$

- Average Direct Treatment Effect

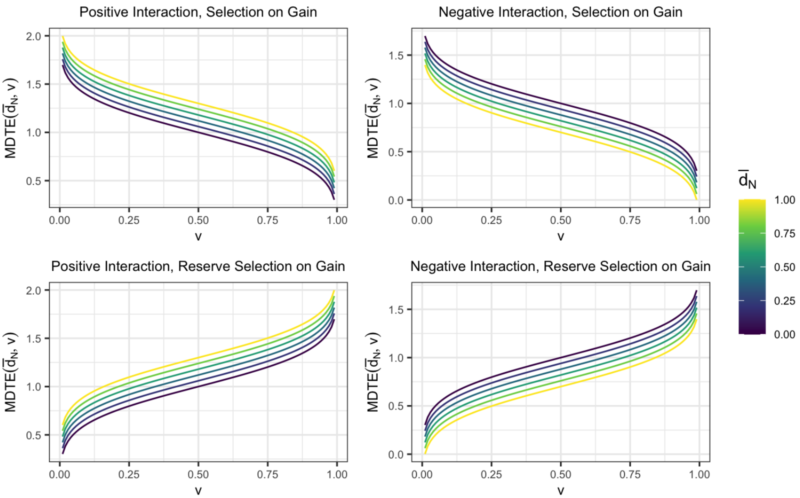
$$\left(\delta^{(1)} - \delta^{(0)}\right) \bar{d}_{\mathcal{N}} + \left(\beta^{(1)} - \beta^{(0)}\right) \mathbb{E}\left[X_i\right]$$

Beyond Average

- Marginal Direct Treatment Effect

$$
\begin{aligned}
MDTE\left(\bar{d}_{\mathcal{N}}, v\right) &:= \mathbb{E}\left[Y_i^{(1)} - Y_i^{(0)} \mid \bar{D}_{\mathcal{N}i} = \bar{d}_{\mathcal{N}}, \overbrace{V_i = v}^{\text{unobservable resistance to treatment}}, X_i\right] \\
&= \left(\delta^{(1)} - \delta^{(0)}\right) \bar{d}_{\mathcal{N}} + \mathbb{E}\left[\epsilon_i^{(1)} - \epsilon_i^{(0)} \mid V_i = v\right] + \left(\beta^{(1)} - \beta^{(0)}\right) \mathbb{E}\left[X_i\right] \\
&= \underbrace{\left(\delta^{(1)} - \delta^{(0)}\right)}_{\text{patterns of interaction effects}} \bar{d}_{\mathcal{N}} - \underbrace{\left(\sigma_{1D} - \sigma_{0D}\right)}_{\text{patterns of selection into treatment}} F^{-1}(v) + \left(\beta^{(1)} - \beta^{(0)}\right) \mathbb{E}\left[X_i\right].
\end{aligned}
$$

Marginal Direct Treatment Effects



Outline

Estimation and Inference

- We develop estimators for parameters $\delta^{(1)}, \delta^{(0)}, \beta^{(1)}, \beta^{(0)}, \sigma_{1D}, \sigma_{0D}$, and in turn for causal estimands.
- Rewrite: $\mathbf{P}_i = [Z_i, X_i]$, $\gamma = \left[\alpha, \beta^{(D)}\right]'$,
  and $\mathbf{Q}_i = [\bar{D}_{\mathcal{N}i}, X_i]'$, $\beta_1 = \left[\delta^{(1)}, \beta^{(1)}\right]'$, $\beta_0 = \left[\delta^{(0)}, \beta^{(0)}\right]'$.
- The model can be more compactly written as

$$
\begin{aligned}
D_i^* &= \mathbf{P}_i\gamma + \epsilon_i^{(D)} \\
Y_i^{(1)} &= \mathbf{Q}_i\beta_1 + \epsilon_i^{(1)} \\
Y_i^{(0)} &= \mathbf{Q}_i\beta_0 + \epsilon_i^{(0)} \\
D_i &= \mathbb{1}\{D_i^* \geq 0\} \\
Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}
\end{aligned}
$$

... similar to the *Generalised Roy Model* (Heckman and Vytlacil, 2007).

Estimation and Inference

♣ Rewrite model in a compact form

$$D_i^* = \mathbf{P}_i\gamma + \epsilon_i^{(D)}$$
$$Y_i^{(1)} = \mathbf{Q}_i\beta_1 + \epsilon_i^{(1)}$$
$$Y_i^{(0)} = \mathbf{Q}_i\beta_0 + \epsilon_i^{(0)}$$

$$D_i = \mathbb{1}\{D_i^* \geq 0\}$$
$$Y_i = D_iY_i^{(1)} + (1 - D_i)Y_i^{(0)}$$

	Y	$Y^{(1)}$	$Y^{(0)}$	D	D^*
obs.1	0.34	1.88	0.34	0	-2.94
obs.2	0.85	0.24	0.85	0	-6.44
obs.3	1.55	1.55	2.15	1	0.63
obs.4	0.50	1.45	0.5	0	-1.12
obs.5	4.01	4.01	3.49	1	3.62
obs.6	3.53	3.53	0.98	1	0.79
obs.7	3.50	3.5	2.49	1	6.62
obs.8	0.81	1.6	0.81	0	-2.66

unobservable!

A Missing Data Perspective?

♣ Simplified Model

$$
\begin{aligned}
D_i^* &= \mathbf{P}_i \gamma + \epsilon_i^{(D)} \\
Y_i^{(1)} &= \mathbf{Q}_i \beta_1 + \epsilon_i^{(1)} \\
Y_i^{(0)} &= \mathbf{Q}_i \beta_0 + \epsilon_i^{(0)}
\end{aligned}
$$

$$
\begin{aligned}
D_i &= \mathbb{1}\{D_i^* \geq 0\} \\
Y_i &= D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}
\end{aligned}
$$

♣ Bayesian Data Augmentation

- *Imputation step (I-step)*: impute the **missing data** given the observed data and model parameters.
- *Posterior step (P-step)*: update the posterior distributions of model parameters given completed data.

Estimation Algorithm

Algorithm 1: Bayesian Data Augmentation

Data: $\{Y_i, D_i, \mathbf{P}_i, \mathbf{Q}_i\}_{i=1}^{n}$

Result: $\hat{\Delta}$, where Δ is parameter of interest.

Procedure

Initialize parameters $s = 0, \theta^{[0]}, \Sigma^{[0]}$ for MCMC-chains

while $s < S$ **do**

  Sample $Y_{miss}^{[s+1]}$ from $p(Y_{miss} \mid Y, D, \theta^{[s]}, \Sigma^{[s]})$

  Sample $D^{*[s+1]}$ from $p(D^* \mid Y, D, \theta^{[s]}, \Sigma^{[s]})$

  Sample $\theta^{[s+1]}$ from $p(\theta \mid Y, D, \Sigma^{[s]}, Y_{miss}^{[s+1]}, D^{*[s+1]})$

  Sample $\Sigma^{[s+1]}$ from $p(\Sigma \mid Y, D, \theta^{[s+1]}, Y_{miss}^{[s+1]}, D^{*[s+1]})$

  Derive $\Delta^{[s+1]}$

end

Calculate $\hat{\Delta} = \frac{1}{S} \sum_{s=\text{nb}+1}^{S} \Delta^{[s]}$

end procedure

Outline

Simulation - DGP

For $i = 1, \ldots, n$ and $k = 1, \ldots, 5$:

$$
\begin{aligned}
Z_i, X_{1i}, \ldots, X_{5i} &\overset{iid}{\sim} \mathcal{N}(0,1), \\
\beta_k^{(D)}, \beta_k^{(1)}, \beta_k^{(0)} &\overset{iid}{\sim} \mathcal{U}(-1,1), \\
D_i &= \mathbb{1}\left\{1.5Z_i + \mathbf{X}_i\beta^{(D)} + \epsilon_i^{(D)} \geq 0\right\}, \\
\bar{D}_{\mathcal{N}i} &= \sum_{j=1, j\neq i}^{N} w_{ij} D_j, \quad \sum_{j=1, j\neq i} w_{ij} = 1, \\
Y_i &= \begin{cases} \delta^{(1)} \bar{D}_{\mathcal{N}i} + 2 + \mathbf{X}_i\beta^{(1)} + \epsilon_i^{(1)} \\ \delta^{(0)} \bar{D}_{\mathcal{N}i} + 1 + \mathbf{X}_i\beta^{(0)} + \epsilon_i^{(0)} \end{cases} .
\end{aligned}
$$

where weight matrix W is block diagonal and row-normalized (Liu and Lee, 2010).

Simulation - Scenarios

1. Sample size $n \in \{100, 500, 1000\}$
2. The presence of spillovers
   - *with spillovers*: $\delta^{(1)} = 1.5$; $\delta^{(0)} = 0.5$
   - *no spillovers*: $\delta^{(1)} = \delta^{(0)} = 0$
3. The joint distribution of the error terms
   - *normality* - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]' \overset{iid}{\sim} \mathcal{N}(0, \Sigma); \quad \Sigma = \begin{bmatrix} 1 & 0.9 & 0.7 \\ & 1 & 0.6 \\ & & 1 \end{bmatrix}$$

   - *non-normality* - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]' \overset{iid}{\sim} \frac{1}{3}\mathcal{N}(0, \Sigma_1) + \frac{2}{3}\mathcal{N}(0, \Sigma_2)$$

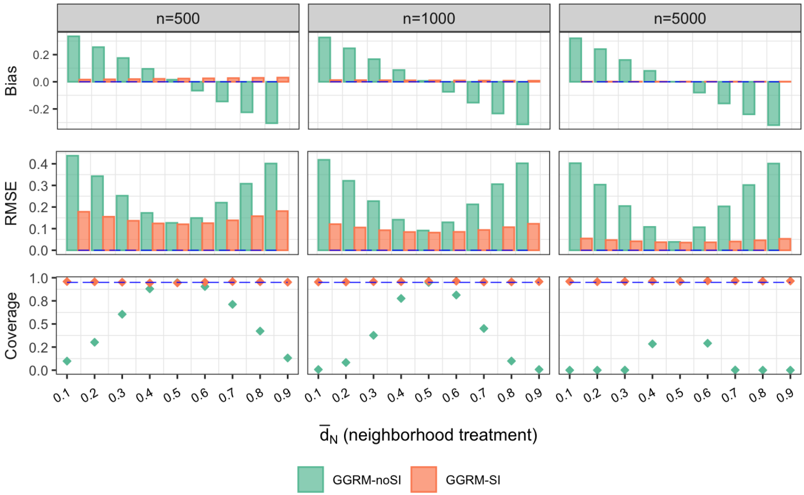
Table 1.
Simulation results across 1000 replicates.

Model	Metric	n	Quantities of Interest $\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)} - \delta^{(0)}$	$\sigma_{1D} - \sigma_{0D}$
		True Value	1.500	0.500	1.000	0.200
GGRM-noSI	Bias	500	-1.500	-0.500	-1.000	-0.035
		1000	-1.500	-0.500	-1.000	-0.021
		5000	-1.500	-0.500	-1.000	-0.004
	RMSE	500	1.500	0.500	1.000	0.134
		1000	1.500	0.500	1.000	0.093
		5000	1.500	0.500	1.000	0.044
	Coverage	500	0.000	0.000	0.000	0.947
		1000	0.000	0.000	0.000	0.947
		5000	0.000	0.000	0.000	0.947
GGRM-SI	Bias	500	-0.001	-0.024	0.023	-0.050
		1000	0.000	0.007	-0.007	-0.028
		5000	0.000	0.000	0.000	-0.006
	RMSE	500	0.226	0.244	0.334	0.133
		1000	0.144	0.171	0.225	0.089
		5000	0.066	0.077	0.100	0.040
	Coverage	500	0.954	0.946	0.954	0.934
		1000	0.968	0.950	0.953	0.947
		5000	0.960	0.951	0.960	0.957

Table 2. Simulation Results for Average Direct Treatment Effects

		n = 500			n = 1000			n = 5000		
Model	Grid	Bias	RMSE	Coverage	Bias	RMSE	Coverage	Bias	RMSE	Coverage
GGRM-noSI	0.1	0.419	0.437	0.098	0.408	0.418	0.007	0.401	0.403	0.000
	0.2	0.319	0.343	0.302	0.308	0.321	0.083	0.301	0.303	0.000
	0.3	0.219	0.252	0.605	0.208	0.227	0.377	0.201	0.204	0.001
	0.4	0.119	0.173	0.880	0.108	0.141	0.776	0.101	0.108	0.284
	0.5	0.019	0.127	0.948	0.008	0.091	0.949	0.001	0.038	0.961
	0.6	-0.081	0.149	0.905	-0.092	0.129	0.813	-0.099	0.106	0.291
	0.7	-0.181	0.220	0.712	-0.192	0.212	0.451	-0.199	0.203	0.002
	0.8	-0.281	0.308	0.423	-0.292	0.306	0.099	-0.299	0.302	0.000
	0.9	-0.381	0.401	0.132	-0.392	0.402	0.007	-0.399	0.401	0.000
GGRM-SI	0.1	0.020	0.178	0.959	0.016	0.121	0.954	0.002	0.054	0.960
	0.2	0.023	0.155	0.956	0.015	0.105	0.954	0.002	0.047	0.957
	0.3	0.025	0.136	0.952	0.014	0.093	0.957	0.002	0.041	0.960
	0.4	0.027	0.124	0.944	0.014	0.084	0.957	0.002	0.037	0.963
	0.5	0.029	0.120	0.944	0.013	0.082	0.962	0.002	0.036	0.961
	0.6	0.032	0.125	0.952	0.012	0.085	0.963	0.002	0.037	0.966
	0.7	0.034	0.138	0.958	0.011	0.094	0.954	0.002	0.040	0.966
	0.8	0.036	0.158	0.956	0.011	0.107	0.955	0.002	0.046	0.963
	0.9	0.039	0.181	0.952	0.010	0.122	0.958	0.002	0.053	0.965

Figure 1. Simulation Results for Average Direct Treatment Effects



Scenario 1' - No Spillover

Table 3. Simulation Results for Scenario 1' (*No spillover*)

Metric	n	$\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)} - \delta^{(0)}$	$\sigma_{1D} - \sigma_{0D}$
		Quantities of Interest			
	True Value	0.000	0.000	0.000	0.200
Bias	500	0.000	-0.024	0.023	-0.050
	1000	0.000	0.007	-0.007	-0.028
	5000	0.000	0.000	0.000	-0.006
RMSE	500	0.227	0.244	0.335	0.132
	1000	0.144	0.171	0.226	0.089
	5000	0.066	0.077	0.100	0.040
Coverage	500	0.952	0.944	0.955	0.937
	1000	0.970	0.951	0.951	0.946
	5000	0.957	0.951	0.956	0.953

Simulation - Results (cont.)

Table 4. Simulation Results for Scenario 2' (*Non-normality*)

Metric	n	$\delta^{(1)}$	$\delta^{(0)}$	$\delta^{(1)} - \delta^{(0)}$	$\sigma_{1D} - \sigma_{0D}$
		Quantities of Interest			
	True Value	1.500	0.500	1.000	0.200
Bias	500	0.008	0.000	0.008	-0.012
	1000	-0.002	-0.001	-0.002	0.005
	5000	0.000	0.002	-0.002	0.025
RMSE	500	0.238	0.264	0.352	0.190
	1000	0.175	0.188	0.260	0.135
	5000	0.076	0.081	0.110	0.072
Coverage	500	0.967	0.956	0.965	0.949
	1000	0.945	0.937	0.938	0.945
	5000	0.962	0.958	0.960	0.904

Simulation - Remarks

- Bayesian estimator performs well in terms of bias, RMSE and coverage rate.
- Inclusion of neighborhood treatment term is plausible, regardless of whether spillovers are present in true DGP.
- Neglecting neighborhood treatment leads to a larger bias and a lower coverage rate, even when the causal estimand is the direct treatment effect.

Outline

Impact of the Opportunity Zones (OZ) tax incentive

♣ **Q:** Direct and indirect effects of *OZ program* on census tracts' *housing unit growth*?

- Part of the Tax Cuts and Jobs Act (Dec 2017): offers investment tax incentives to designated census tracts nationwide to promote economic development in distressed communities.
- Selection process:

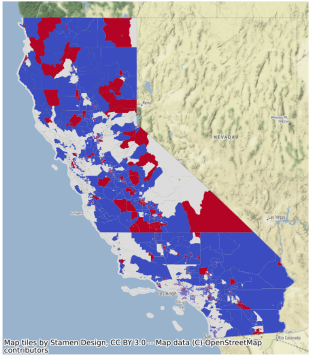
  eligibility: poverty rates above 20% or median family incomes below 80% of the area median income.
  nomination: each state governor had 90-120 days to nominate 25% of their eligible tracts.
  approval: on June 14, 2018, U.S. Treasury and IRS certified a list of Qualified Opportunity Zones (QOZs).

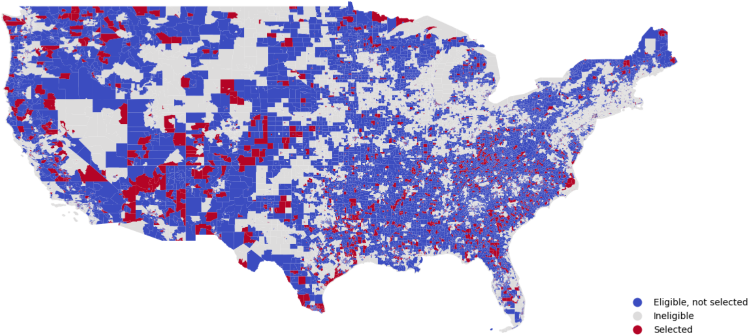
- Mixed evidence regarding impact on targeted areas and potential spillovers:

  Corinth and Feldman (2024), Freedman et al. (2023), Chen et al. (2023), and Wheeler (2022), etc.

OZ status mapping

California 06





Model Specification

- Individual Treatment - being selected into the OZ program

$$QOZ_i = \mathbb{1}\left\{\alpha Political_i + \beta_0^{(D)} + \sum_{k=1}^{K} \beta_k^{(D)} Demographic_{k,i} + \epsilon_i^{(D)} > 0\right\}$$

- Neighborhood Treatment

$$\overline{QOZ}_i = \sum_{j=1, j\neq i} w_{ij} QOZ_j; \quad \sum_{j=1, j\neq i} w_{ij} = 1$$

- Revealed Outcome

$$\%\Delta Housing_i = \begin{cases} \delta^{(1)}\overline{QOZ}_i + \beta_0^{(1)} + \sum_{k=1}^{K} \beta_k^{(1)} Demographic_{k,i} + \epsilon_i^{(1)}, & \text{if} \quad QOZ_i = 1 \\ \delta^{(0)}\overline{QOZ}_i + \beta_0^{(0)} + \sum_{k=1}^{K} \beta_k^{(0)} Demographic_{k,i} + \epsilon_i^{(0)}, & \text{if} \quad QOZ_i = 0 \end{cases}$$

Eligible Census Tracts in California

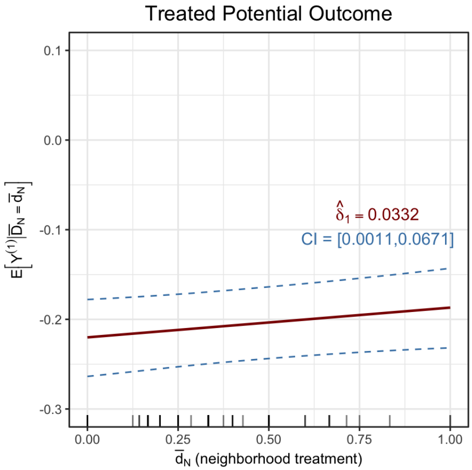
Table 4. Summary Statistics and Balancing Tests

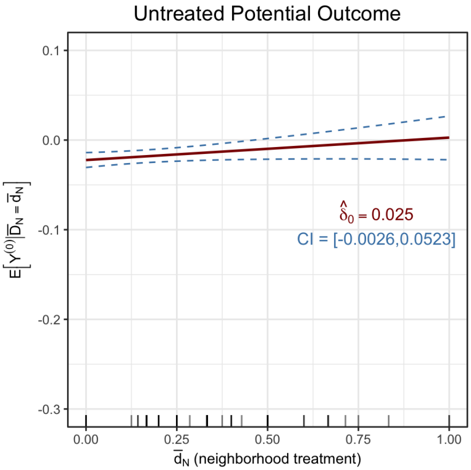
Variables	All tracts (n=3699)	QOZs (n=727)	Non-QOZs (n=2972)	QOZs – Non-QOZs		
	Mean (std)	Mean (std)	Mean (std)	Diff.Mean		t-statistic
Outcome						
Housing Unit Growth	0.03 (0.17)	0.04 (0.14)	0.03 (0.17)	0.02	*	2.57
Observed Characteristics						
Political Affiliation	0.79 (0.41)	0.82 (0.38)	0.78 (0.42)	0.04	**	2.72
Poverty Rate	0.19 (0.09)	0.27 (0.09)	0.17 (0.08)	0.09	***	24.64
Median Earnings	10.17 (0.31)	10.01 (0.26)	10.21 (0.30)	-0.21	***	-18.60
Employment Rate	0.29 (0.07)	0.26 (0.07)	0.29 (0.07)	-0.04	***	-12.87
% White	0.56 (0.21)	0.53 (0.20)	0.56 (0.21)	-0.04	***	-4.35
% Native	0.90 (0.04)	0.89 (0.04)	0.91 (0.04)	-0.02	***	-9.05
% Higher ed.	0.15 (0.09)	0.11 (0.07)	0.16 (0.09)	-0.05	***	-15.23
% Rent	0.57 (0.21)	0.67 (0.19)	0.54 (0.21)	0.13	***	16.47
Population	4509.55 (1613.93)	4305.31 (1476.18)	4559.51 (1642.24)	-254.20	***	-4.07

Table 4.
Estimation Results

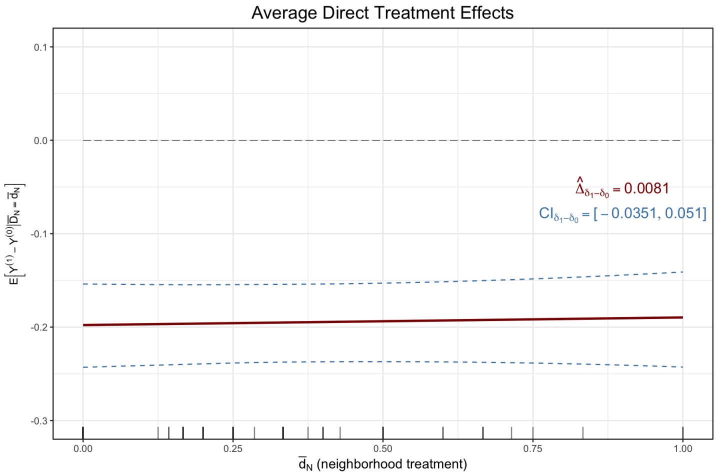
	Mean	Std	LB	UB
Treatment Decision Equation				
Political Affiliation (α)	0.116	0.053	0.017	0.221
Intercept ($\beta_0^{(D)}$)	1.184	1.141	-1.077	3.410
Poverty Rate ($\beta_1^{(D)}$)	4.951	0.334	4.307	5.612
Median Income ($\beta_2^{(D)}$)	-0.317	0.113	-0.538	-0.095
Employment Rate ($\beta_3^{(D)}$)	0.530	0.459	-0.350	1.423
Outcome Equation for QOZs				
Neighborhood Treatment ($\delta^{(1)}$)	0.033	0.017	0.001	0.067
Intercept ($\beta_0^{(1)}$)	-0.573	0.295	-1.164	0.001
Poverty Rate ($\beta_1^{(1)}$)	0.857	0.101	0.667	1.054
Median Income ($\beta_2^{(1)}$)	0.017	0.030	-0.041	0.074
Employment Rate ($\beta_3^{(1)}$)	0.057	0.109	-0.153	0.270
Outcome Equation for Non-QOZs				
Neighborhood Treatment ($\delta^{(0)}$)	0.025	0.014	-0.003	0.052
Intercept ($\beta_0^{(0)}$)	-0.718	0.152	-1.016	-0.420
Poverty Rate ($\beta_1^{(0)}$)	-0.308	0.050	-0.408	-0.212
Median Income ($\beta_2^{(0)}$)	0.080	0.015	0.050	0.109
Employment Rate ($\beta_3^{(0)}$)	-0.200	0.059	-0.315	-0.084
Pattern of Interaction Effect				
$\Delta_{\delta^{(1)}-\delta^{(0)}}$	0.008	0.022	-0.035	0.051
Pattern of Selection into Treatment				
$\Delta_{\sigma_{1D}-\sigma_{0D}}$	0.329	0.015	0.301	0.358

Indirect Treatment Effects



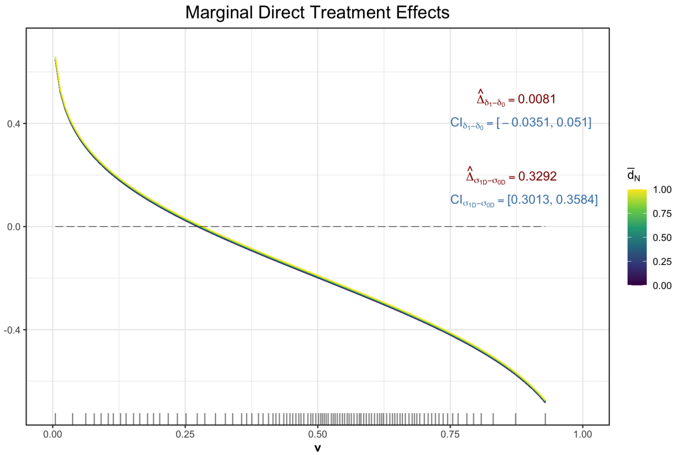


Average Direct Treatment Effects



$\bar{d}_{\mathcal{N}}$	Mean (std)	CI95
0.1	-0.1970 (0.0223)	[-0.2413, -0.1546]
0.2	-0.1962 (0.0219)	[-0.2394, -0.1546]
0.3	-0.1954 (0.0218)	[-0.2377, -0.1544]
0.4	-0.1945 (0.0218)	[-0.2373, -0.1538]
0.5	-0.1937 (0.0221)	[-0.2373, -0.1532]
0.6	-0.1929 (0.0225)	[-0.2373, -0.1514]
0.7	-0.1921 (0.0232)	[-0.2379, -0.1493]
0.8	-0.1913 (0.0241)	[-0.2387, -0.1471]
0.9	-0.1905 (0.0251)	[-0.2406, -0.1441]

Marginal Direct Treatment Effects



v	Mean (std)	CI95
0.1	0.2258 (0.0127)	[0.2010, 0.2506]
0.2	0.0809 (0.0137)	[0.0538, 0.1075]
0.3	-0.0235 (0.0162)	[-0.0554, 0.0071]
0.4	-0.1127 (0.0190)	[-0.1502, -0.0766]
0.5	-0.1961 (0.0219)	[-0.2393, -0.1546]
0.6	-0.2796 (0.0251)	[-0.3288, -0.2320]
0.7	-0.3688 (0.0286)	[-0.4247, -0.3151]
0.8	-0.4732 (0.0329)	[-0.5376, -0.4115]
0.9	-0.6181 (0.0391)	[-0.6945, -0.5451]

By Sub-populations

Table 4. Summary of Direct Treatment Effects

ADTE		ADTT		ADTUT	
Mean (std)	CI95	Mean (std)	CI95	Mean (std)	CI95
-0.196 (0.022)	[-0.239, -0.155]	0.041 (0.015)	[0.012, 0.069]	-0.298 (0.026)	[-0.349, -0.249]

- Eligible but unselected tracts (non-QOZs) continue to face disadvantages: No positive spillover effects found, and expanding the OZ tax credit to these communities would not be effective.

Outline

Conclusion

1. Many policies we care about have Endogenous Selection into Treatment and Spillovers.
   - need to be careful when estimating effects since certain restrictions are necessary to identify causal estimands.

2. My approach explicitly models endogenous selection into treatment and employs spatial/network data to capture spillovers in the form of a neighborhood treatment term.
   - allows for heterogeneous direct effects across individuals and general interference.
   - embraces Bayesian methods to relax parametric assumptions (*further steps*).

Thank you!

Reference I

P. M. Aronow and C. Samii. Estimating average causal effects under general interference, with application to a social network experiment. 2017.

J. Chen, E. Glaeser, and D. Wessel. Jue insight: The (non-) effect of opportunity zones on housing prices. Journal of Urban Economics, 133:103451, 2023.

K. Corinth and N. Feldman. Are opportunity zones an effective place-based policy? Journal of Economic Perspectives, 38(3):113–136, 2024.

P. Ding. Bayesian robust inference of sample selection using selection-t models. Journal of Multivariate Analysis, 124:451–464, 2014.

F. J. DiTraglia, C. Garcia-Jimeno, R. OKeeffe-ODonovan, and A. Sanchez-Becerra. Identifying causal effects in experiments with spillovers and non-compliance. Journal of Econometrics, 2023.

O. Doğan and S. Taşpinar. Bayesian inference in spatial sample selection models. Oxford Bulletin of Economics and Statistics, 80(1):90–121, 2018.

Reference II

L. Forastiere, E. M. Airoldi, and F. Mealli. Identification and estimation of treatment and interference effects in observational studies on networks. Journal of the American Statistical Association, 116 (534):901–918, 2021.

L. Forastiere, F. Mealli, A. Wu, and E. M. Airoldi. Estimating causal effects under network interference with bayesian generalized propensity scores. Journal of Machine Learning Research, 23(289):1–61, 2022.

M. Freedman, S. Khanna, and D. Neumark. Jue insight: The impacts of opportunity zones on zone residents. Journal of Urban Economics, 133:103407, 2023.

J. J. Heckman and E. J. Vytlacil. Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. Handbook of econometrics, 6: 4875–5143, 2007.

Reference III

J. J. Heckman, H. F. Lopes, and R. Piatek. Treatment effects: A bayesian perspective. Econometric reviews, 33(1-4):36–67, 2014.

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

M. P. Leung. Treatment and spillover effects under network interference. Review of Economics and Statistics, 102(2):368–380, 2020.

M. P. Leung and P. Loupos. Unconfoundedness with network interference. arXiv preprint arXiv:2211.07823, 2022.

M. Li, D. J. Poirier, and J. L. Tobias. Do dropouts suffer from dropping out? estimation and prediction of outcome gains in generalized selection models. Journal of Applied Econometrics, 19(2): 203–225, 2004.

X. Liu and L.-f. Lee. Gmm estimation of social interaction models with centrality. Journal of Econometrics, 159(1):99–115, 2010.

Reference IV

C. F. Manski. Identification of treatment response with social interactions. The Econometrics Journal, 16(1):S1–S23, 2013.

E. L. Ogburn, O. Sofrygin, I. Diaz, and M. J. Van der Laan. Causal inference for social network data. Journal of the American Statistical Association, pages 1–15, 2022.

A. Sanchez-Becerra. The network propensity score: Spillovers, homophily, and selection into treatment. arXiv preprint arXiv:2209.14391, 2022.

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

P. Toulis and E. Kao. Estimation of causal peer influence effects. In International conference on machine learning, pages 1489–1497. PMLR, 2013.

M. J. van der Laan and O. Sofrygin. Semi-parametric estimation and inference for the mean outcome of the single time-point intervention in a causally connected population. 2017.

Reference V

H. Wheeler. Locally optimal place-based policies: Evidence from opportunity zones. Unpublished working paper. https://hbwheeler. github. io/files/JMP_HW. pdf, 2022.

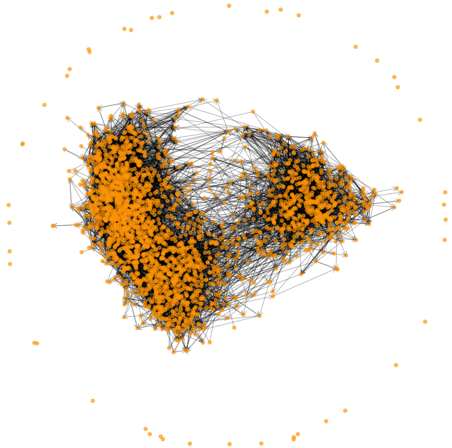
Y. Yuan and K. M. Altenburger. A two-part machine learning approach to characterizing network interference in a/b testing, 2022.

Empirical Monte Carlo Study

- I employ Add Health friendship network data and mimic an evaluation of *Social-Emotional Learning (SEL)-Focused After-School Programs* on youth's prosocial development.
  - actual network structure defines the spillover patterns;
  - real covariates are considered as observed characteristics;
  - treatment and outcome are generated using different data generating processes to analyse performance of the estimation procedures under different scenarios.

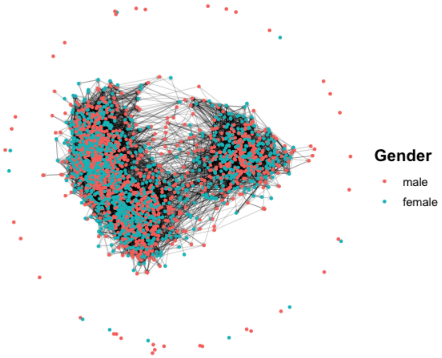
Add Health friendship network data

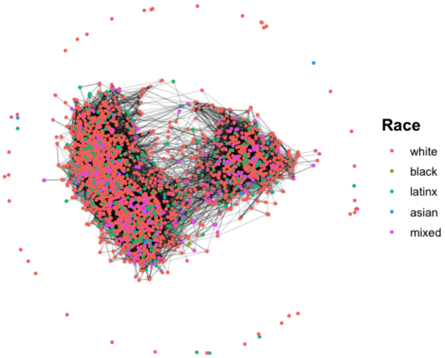
- A nationally representative longitudinal study of adolescents in grades 7–12 in the US between September 1994 and April 1995.
- The largest community: $n = 2534$ students.
- Each node represents a student, and network links are measured using student nomination (*i.e., their best friends, up to 5 females and up to 5 males*).



Add Health friendship network data

Add Health friendship network data





Data Generating Process

- Individual Treatment - enrolling in the SEL-focus after-school program

$$D_i = \mathbb{1}\left\{1.5Z_i - 0.2X_{gender,i} - X_{grade,i} + X_{race,i} + \epsilon_i^{(D)} \geq 0\right\}$$

- Neighbourhood Treatment

$$\bar{D}_i = \sum_{j=1, j\neq i} w_{ij} D_j; \quad \sum_{j=1, j\neq i} w_{ij} = 1$$

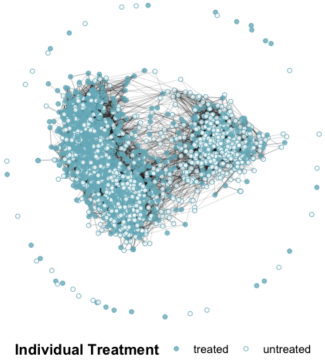
- Treated and Untreated Potential Outcomes

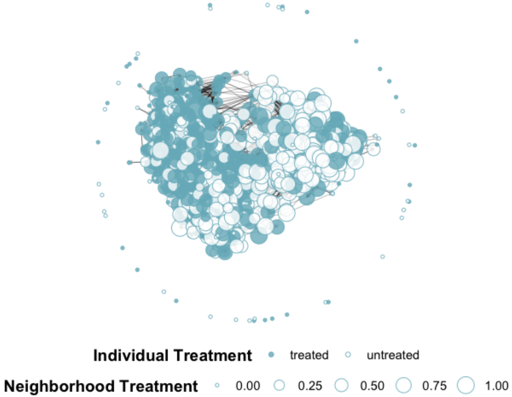
$$Y_i^{(1)} = \delta^{(1)}\bar{D}_i + 2 - 0.5X_{gender,i} + 0.3X_{grade,i} + 0.2X_{race,i} + \epsilon_i^{(1)}$$
$$Y_i^{(0)} = \delta^{(0)}\bar{D}_i + 1 + 0.3X_{gender,i} - 0.4X_{grade,i} + 0.1X_{race,i} + \epsilon_i^{(0)}$$

- Revealed Outcome

$$Y_i = D_i Y_i^{(1)} + (1 - D_i) Y_i^{(0)}$$

Treatment status - what if my friend is treated?





$[2 \times 2]$ scenarios

1. The presence of spillovers
   - *without spillovers*: $\delta^{(1)} = \delta^{(0)} = 0$
   - *with spillovers*: $\delta^{(1)} = 1.5;\quad \delta^{(0)} = 0.5$
2. The joint distribution of the error terms
   - *a normal distribution* - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]' \overset{ind}{\sim} \mathcal{N}(0, \Sigma); \quad \Sigma = \begin{bmatrix} 1 & 0.9 & 0.7 \\ & 1 & 0.6 \\ & & 1 \end{bmatrix}$$

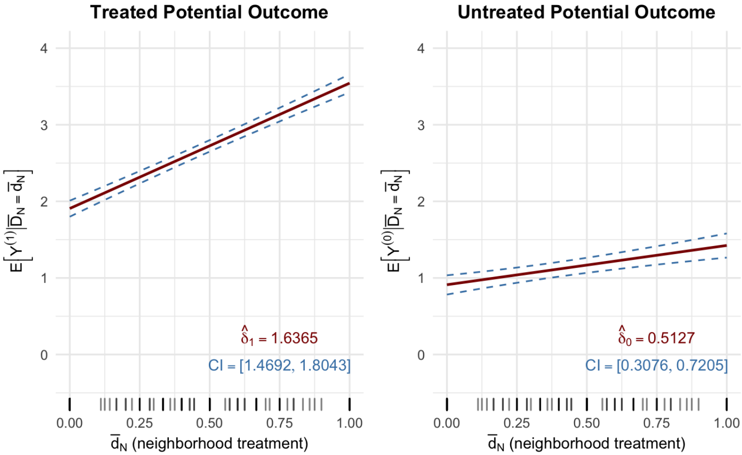
   - *a finite mixture of normal distribution* - for $i = 1, \ldots, n$

$$\epsilon_i = \left[\epsilon_i^{(D)}, \epsilon_i^{(1)}, \epsilon_i^{(0)}\right]' \overset{ind}{\sim} \frac{1}{3}\mathcal{N}(0, \Sigma_1) + \frac{2}{3}\mathcal{N}(0, \Sigma_2)$$

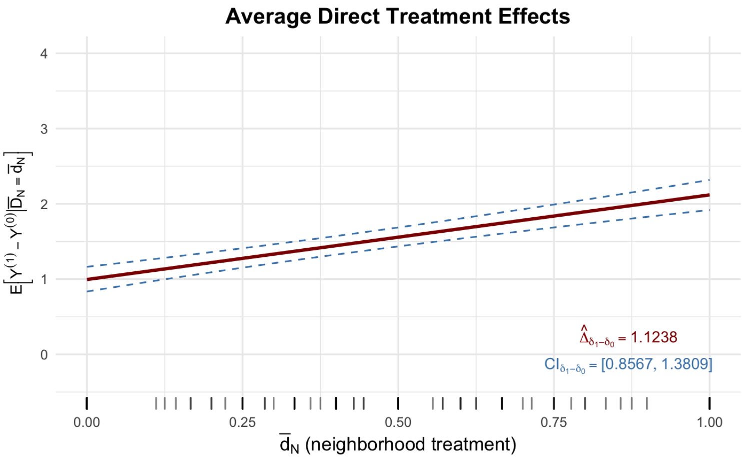
♣ **Simulation results**:

- Bayesian estimator performs well in terms of bias, RMSE and coverage rate.
- Inclusion of neighbourhood treatment term is plausible, regardless of whether spillovers are present

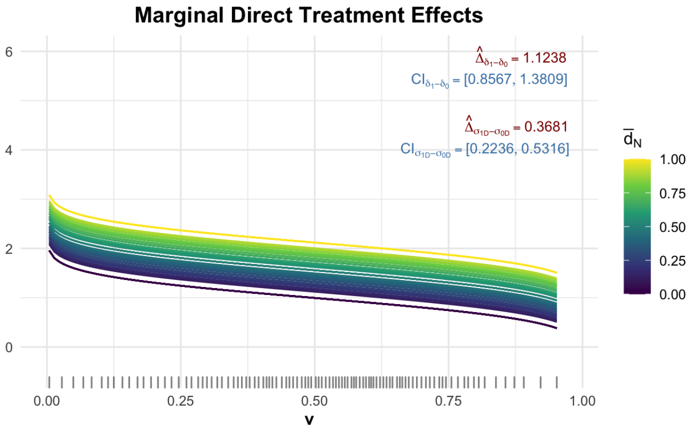
Spillovers exist: $\delta^{(1)} = 1.5; \delta^{(0)} = 0.5$



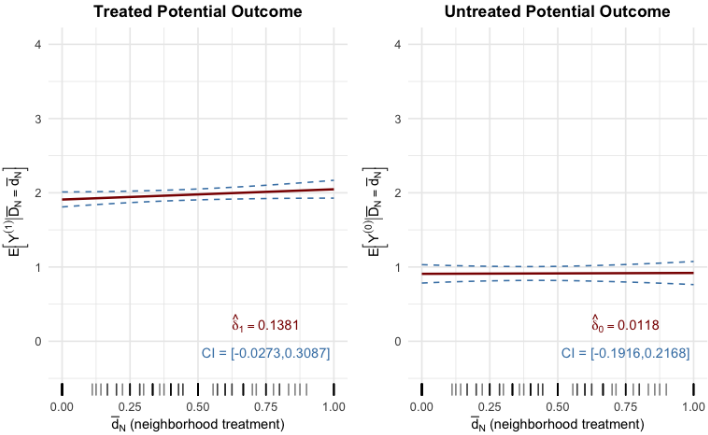
Spillovers exist: $\delta^{(1)} = 1.5; \delta^{(0)} = 0.5$



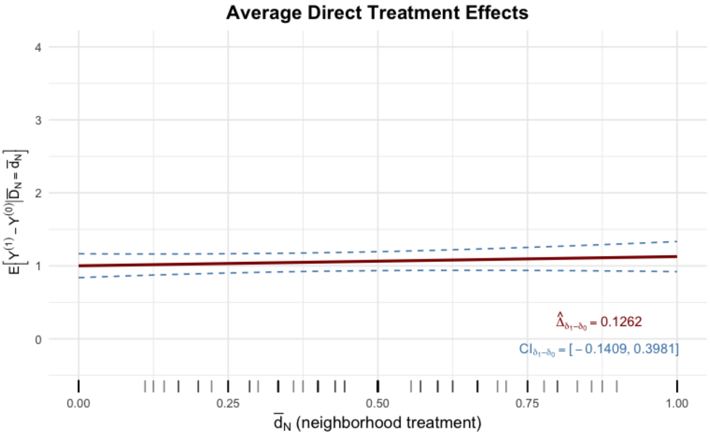
Spillovers exist: $\delta^{(1)} = 1.5; \delta^{(0)} = 0.5$



No Spillovers: $\delta^{(1)} = \delta^{(0)} = 0$



No Spillovers: $\delta^{(1)} = \delta^{(0)} = 0$



No Spillovers: $\delta^{(1)} = \delta^{(0)} = 0$