Estimating within-cluster and between-cluster spillover effects in randomized saturation designs

Sizhu Lu

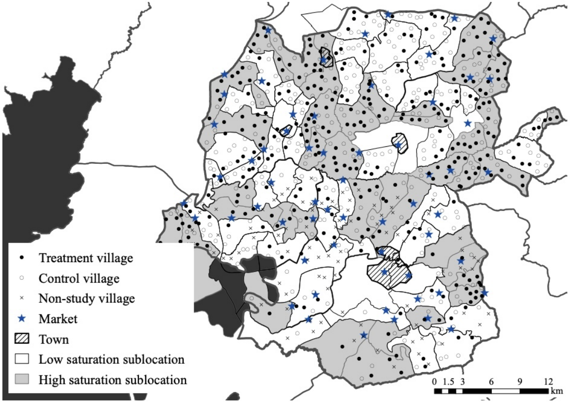
Casual Causal Group Meeting
September 4, 2025

Joint work with Lei Shi and Peng Ding

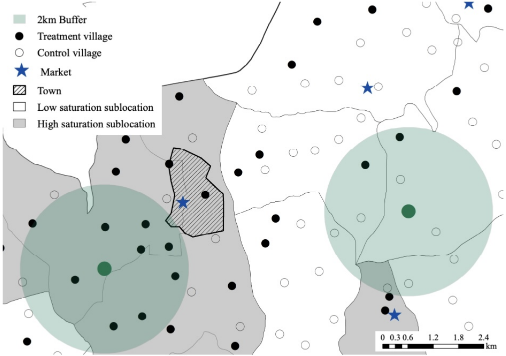
Outline

Motivating example: large-scale cash transfer experiment in rural Kenya (Egger et al., 2022)

- Two-stage *randomized saturation* design:
  - **Stage 1:** Sublocations randomized to high and low saturation.
  - **Stage 2:** Villages randomized to treatment at the assigned saturation rate.
- Geography (markets, labor, mobility) induces interference *within* and *between* sublocations.



Spatial exposure to treatment



- Evidence of meaningful spillovers on non-recipient households and local markets.
- Positive effects on local enterprise activity and consumption reported.

Randomized saturation designs

- Randomize treatment in two stages (Hudgens and Halloran, 2008; Baird et al., 2018; Basse and Feller, 2018):
  1. Stage 1 (cluster level): Randomize clusters (communities, schools, villages, networks) to different saturation levels (high vs. low coverage).
  2. Stage 2 (unit level): Randomize individual units within each cluster to treatment, according to the assigned saturation rate.
- Enables identification of both direct and spillover effects.

Randomized saturation designs

- Randomize treatment in two stages (Hudgens and Halloran, 2008; Baird et al., 2018; Basse and Feller, 2018):
  1. Stage 1 (cluster level): Randomize clusters (communities, schools, villages, networks) to different saturation levels (high vs. low coverage).
  2. Stage 2 (unit level): Randomize individual units within each cluster to treatment, according to the assigned saturation rate.
- Enables identification of both direct and spillover effects.
- Widely-used in public health and social sciences: deworming on health (Miguel and Kremer, 2004), vaccination on health (Ali et al., 2005), labor market policies on displacement (Crépon et al., 2013), school governance on learning (Pradhan et al., 2014), microcredit on business outcomes (Banerjee et al., 2015), education on fertility (Duflo et al., 2015), cash transfers on school attendance (Barrera-Osorio et al., 2019).

Motivation

- Common assumption in prior work: spillovers **only** occur within clusters, no interference between clusters.

Motivation

- Common assumption in prior work: spillovers **only** occur within clusters, no interference between clusters.
- Not reasonable in many cases:
  - Real-world clusters are rarely isolated.
  - Networks, mobility, markets, or environmental channels can connect units across cluster boundaries.
  - Ignoring between-cluster interference can bias effect estimates and mislead policy conclusions.

Our contribution and roadmap

- Formalize analysis in randomized saturation designs, allowing for *both* within- and between-cluster spillovers.

Our contribution and roadmap

- Formalize analysis in randomized saturation designs, allowing for *both* within- and between-cluster spillovers.
- Estimands: Direct effect, within-cluster indirect effect, and between-cluster indirect effect.

Our contribution and roadmap

- Formalize analysis in randomized saturation designs, allowing for *both* within- and between-cluster spillovers.
- Estimands: Direct effect, within-cluster indirect effect, and between-cluster indirect effect.
- Estimators: Design-based Horvitz–Thompson and Hájek estimators by inverse probability weighting.

Our contribution and roadmap

- Formalize analysis in randomized saturation designs, allowing for *both* within- and between-cluster spillovers.
- Estimands: Direct effect, within-cluster indirect effect, and between-cluster indirect effect.
- Estimators: Design-based Horvitz–Thompson and Hájek estimators by inverse probability weighting.
- Inference: Closed-form asymptotic variance, conservative variance estimators, CLTs, and Wald CIs.

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.
- Exposure mapping (A_i, S_i, H_i): reduces the high-dimensional dependence to a low-dimensional summary, where

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.
- Exposure mapping (A_i, S_i, H_i): reduces the high-dimensional dependence to a low-dimensional summary, where
  - S_i: within-cluster summary, e.g., whether $> 1/2$ villages are getting treated within the same sublocation as i,

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.
- Exposure mapping (A_i, S_i, H_i): reduces the high-dimensional dependence to a low-dimensional summary, where
  - S_i: within-cluster summary, e.g., whether $> 1/2$ villages are getting treated within the same sublocation as i,
  - H_i: between-cluster summary, e.g., whether $> 1/2$ villages are treated among those geographically close to i but located in different sublocations.

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.
- Exposure mapping (A_i, S_i, H_i): reduces the high-dimensional dependence to a low-dimensional summary, where
  - S_i: within-cluster summary, e.g., whether $> 1/2$ villages are getting treated within the same sublocation as i,
  - H_i: between-cluster summary, e.g., whether $> 1/2$ villages are treated among those geographically close to i but located in different sublocations.
- Potential outcome reduces to $Y_i(a, s, h)$.

Notation and exposure mapping

- i: unit/village index.
- $A_i \in \{0, 1\}$: treatment indicator.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$: 2^n potential outcomes.
- Exposure mapping (A_i, S_i, H_i): reduces the high-dimensional dependence to a low-dimensional summary, where
  - S_i: within-cluster summary, e.g., whether $> 1/2$ villages are getting treated within the same sublocation as i,
  - H_i: between-cluster summary, e.g., whether $> 1/2$ villages are treated among those geographically close to i but located in different sublocations.
- Potential outcome reduces to $Y_i(a, s, h)$.
- Finite population regime: fixed potential outcomes, randomness from treatment.

Outline

Controlled direct and indirect effects

- The controlled direct effect of treatment A_i on the outcome, while holding the exposure variables (S_i, H_i) fixed at values (s, h):

$$\mathrm{DE}(s, h) = n^{-1} \sum_{i=1}^{n} Y_i(1, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s, h).$$

Controlled direct and indirect effects

- The controlled direct effect of treatment A_i on the outcome, while holding the exposure variables (S_i, H_i) fixed at values (s, h):

$$\mathrm{DE}(s, h) = n^{-1} \sum_{i=1}^{n} Y_i(1, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s, h).$$

- The within-cluster controlled indirect effect of S_i on the outcome, holding H_i fixed at h:

$$\mathrm{WIE}(s, s', h) = n^{-1} \sum_{i=1}^{n} Y_i(0, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s', h).$$

Controlled direct and indirect effects

- The controlled direct effect of treatment A_i on the outcome, while holding the exposure variables (S_i, H_i) fixed at values (s, h):

$$\mathrm{DE}(s, h) = n^{-1} \sum_{i=1}^{n} Y_i(1, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s, h).$$

- The within-cluster controlled indirect effect of S_i on the outcome, holding H_i fixed at h:

$$\mathrm{WIE}(s, s', h) = n^{-1} \sum_{i=1}^{n} Y_i(0, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s', h).$$

- The between-cluster controlled indirect effect of H_i on the outcome, holding S_i fixed at s:

$$\mathrm{BIE}(s, h, h') = n^{-1} \sum_{i=1}^{n} Y_i(0, s, h) - n^{-1} \sum_{i=1}^{n} Y_i(0, s, h').$$

Marginal direct and indirect effects

- The marginal direct effect of treatment A_i, marginalizing over the distribution of treatments for all other villages:

$$\mathrm{DE} = n^{-1} \sum_{i=1}^{n} \underbrace{E_{\boldsymbol{A}_{(-i)}} \{Y_i(1, S_i, H_i)\}}_{\text{distribution of } \boldsymbol{A}_{(-i)}} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)}} \{Y_i(0, S_i, H_i)\}.$$

Marginal direct and indirect effects

- The marginal direct effect of treatment A_i, marginalizing over the distribution of treatments for all other villages:

$$\mathrm{DE} = n^{-1} \sum_{i=1}^{n} \underbrace{E_{\boldsymbol{A}_{(-i)}} \{ Y_i(1, S_i, H_i) \}}_{\text{distribution of } \boldsymbol{A}_{(-i)}} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)}} \{ Y_i(0, S_i, H_i) \} .$$

- The within-cluster marginal indirect effect:

$$\mathrm{WIE}(s, s') = n^{-1} \sum_{i=1}^{n} \underbrace{E_{\boldsymbol{A}_{(-i)} | S_i = s} \{ Y_i(0, s, H_i) \}}_{\text{conditional distribution } \boldsymbol{A}_{(-i)} \text{ given } S_i = s} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)} | S_i = s'} \{ Y_i(0, s', H_i) \} .$$

Marginal direct and indirect effects

- The marginal direct effect of treatment A_i, marginalizing over the distribution of treatments for all other villages:

$$\mathrm{DE} = n^{-1} \sum_{i=1}^{n} \underbrace{E_{\boldsymbol{A}_{(-i)}} \{Y_i(1, S_i, H_i)\}}_{\text{distribution of } \boldsymbol{A}_{(-i)}} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)}} \{Y_i(0, S_i, H_i)\}.$$

- The within-cluster marginal indirect effect:

$$\mathrm{WIE}(s, s') = n^{-1} \sum_{i=1}^{n} \underbrace{E_{\boldsymbol{A}_{(-i)} \mid S_i = s} \{Y_i(0, s, H_i)\}}_{\text{conditional distribution } \boldsymbol{A}_{(-i)} \text{ given } S_i = s} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)} \mid S_i = s'} \{Y_i(0, s', H_i)\}.$$

- The between-cluster marginal indirect effect:

$$\mathrm{BIE}(h, h') = n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)} \mid H_i = h} \{Y_i(0, S_i, h)\} - n^{-1} \sum_{i=1}^{n} E_{\boldsymbol{A}_{(-i)} \mid H_i = h'} \{Y_i(0, S_i, h')\}.$$

Policy-specific direct and indirect effects

- For a given treatment policy ψ, define the policy-specific direct and indirect effects:

$$\mathrm{DE}_{\psi} = n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(1,S_i,H_i)\} - n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(0,S_i,H_i)\},$$
$$\mathrm{WIE}_{\psi} = n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(0,1,H_i)\} - n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(0,0,H_i)\},$$
$$\mathrm{BIE}_{\psi} = n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(0,S_i,1)\} - n^{-1}\sum_{i=1}^{n} E_{\psi}\{Y_i(0,S_i,0)\},$$

  - subscript ψ: expectation under distribution of (S_i, H_i) conditional on $A_i = a$ induced by policy ψ for $a = 0, 1$.

- Comparison between different policies ψ_1 vs ψ_2, e.g., $\mathrm{DE}_{\psi_1} - \mathrm{DE}_{\psi_2}$.

Outline

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1} \sum_{i=1}^{n} Y_i(a, s, h)$:

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1}\sum_{i=1}^{n} Y_i(a,s,h)$:
  - Define $\mathbb{I}_i(a,s,h) = \mathbb{1}\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a,s,h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1}\sum_{i=1}^{n} Y_i(a,s,h)$:
  - Define $\mathbb{I}_i(a,s,h) = \mathbb{1}\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a,s,h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.
  - Horvitz–Thompson estimator:

$$\hat{Y}^{\mathrm{ht}}(a,s,h) = n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i.$$

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1} \sum_{i=1}^{n} Y_i(a, s, h)$:
  - Define $\mathbb{I}_i(a, s, h) = \mathbb{1}\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a, s, h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.
  - Horvitz–Thompson estimator:

$$\hat{Y}^{\mathrm{ht}}(a, s, h) = n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)}{\pi_i(a, s, h)} Y_i.$$

  - Hájek estimator:

$$\hat{Y}^{\mathrm{haj}}(a, s, h) = n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)}{\pi_i(a, s, h)} Y_i \Big/ n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)}{\pi_i(a, s, h)}.$$

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1}\sum_{i=1}^{n} Y_i(a,s,h)$:
  - Define $\mathbb{I}_i(a,s,h) = \mathbb{1}\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a,s,h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.
  - Horvitz–Thompson estimator:

$$\hat{Y}^{\mathrm{ht}}(a,s,h) = n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i.$$

  - Hájek estimator:

$$\hat{Y}^{\mathrm{haj}}(a,s,h) = n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i \Big/ n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)}.$$

  - Propensity scores can be **simulated** from the design and the network structure.

Estimators for controlled effects

- Estimate each average potential outcome $n^{-1}\sum_{i=1}^{n} Y_i(a,s,h)$:
  - Define $\mathbb{I}_i(a,s,h) = \mathbb{1}\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a,s,h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.
  - Horvitz–Thompson estimator:

$$\hat{Y}^{\mathrm{ht}}(a,s,h) = n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i.$$

  - Hájek estimator:

$$\hat{Y}^{\mathrm{haj}}(a,s,h) = n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i \Big/ n^{-1}\sum_{i=1}^{n} \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)}.$$

  - Propensity scores can be **simulated** from the design and the network structure.
- Controlled effect can be estimated by differences between these average estimators.

Estimators for policy-specific effects

- Need to reweight to target a different policy.

Estimators for policy-specific effects

- Need to reweight to target a different policy.
- Let $\Gamma = \{\gamma_i(a, s, h)\}$ be a set of prespecified unit weights to target different effects.

Estimators for policy-specific effects

- Need to reweight to target a different policy.
- Let $\Gamma = \{\gamma_i(a, s, h)\}$ be a set of prespecified unit weights to target different effects.
- Horvitz–Thompson estimators:

$$\hat{Y}^{\text{ht}}(a, s, h; \Gamma) \quad = \quad n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)\gamma_i(a, s, h)}{\pi_i(a, s, h)} Y_i.$$

Estimators for policy-specific effects

- Need to reweight to target a different policy.
- Let $\Gamma = \{\gamma_i(a, s, h)\}$ be a set of prespecified unit weights to target different effects.
- Horvitz–Thompson estimators:

$$\hat{Y}^{\text{ht}}(a, s, h; \Gamma) \quad = \quad n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)\gamma_i(a, s, h)}{\pi_i(a, s, h)} Y_i.$$

- Hájek-type estimator

$$\hat{Y}^{\text{haj}}(a, s, h) = n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h) \cdot n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)\gamma_i(a, s, h)}{\pi_i(a, s, h)} Y_i \Big/ n^{-1} \sum_{i=1}^{n} \frac{\mathbb{I}_i(a, s, h)\gamma_i(a, s, h)}{\pi_i(a, s, h)}.$$

Choice of weights

- Choose weights to target policy ψ for different effects:

$$\gamma_{i,\psi}^{\mathrm{DE}}(a, s, h) = \pi_{i,\psi}(\cdot, s, h), \quad \gamma_{i,\psi}^{\mathrm{WIE}}(a, s, h) = \pi_{i,\psi}(\cdot, \cdot, h), \quad \gamma_{i,\psi}^{\mathrm{BIE}}(a, s, h) = \pi_{i,\psi}(\cdot, s, \cdot).$$

Choice of weights

- Choose weights to target policy ψ for different effects:

$$\gamma^{\mathrm{DE}}_{i,\psi}(a,s,h) = \pi_{i,\psi}(\cdot,s,h), \quad \gamma^{\mathrm{WIE}}_{i,\psi}(a,s,h) = \pi_{i,\psi}(\cdot,\cdot,h), \quad \gamma^{\mathrm{BIE}}_{i,\psi}(a,s,h) = \pi_{i,\psi}(\cdot,s,\cdot).$$

- All of the above weights can be simulated based on the design.

Choice of weights

- Choose weights to target policy ψ for different effects:

$$\gamma_{i,\psi}^{\mathrm{DE}}(a,s,h)=\pi_{i,\psi}(\cdot,s,h),\quad \gamma_{i,\psi}^{\mathrm{WIE}}(a,s,h)=\pi_{i,\psi}(\cdot,\cdot,h),\quad \gamma_{i,\psi}^{\mathrm{BIE}}(a,s,h)=\pi_{i,\psi}(\cdot,s,\cdot).$$

- All of the above weights can be simulated based on the design.
- Can be used to construct estimators for the policy-specific effects: for $* \in \{\mathrm{ht}, \mathrm{haj}\}$,

$$\begin{aligned}
\hat{\mathrm{DE}}_{\psi}^{*} &= \sum_{s,h=0,1}\{\hat{Y}^{*}(1,s,h;\Gamma_{\psi}^{\mathrm{DE}})-\hat{Y}^{*}(0,s,h;\Gamma_{\psi}^{\mathrm{DE}})\},\\
\hat{\mathrm{WIE}}_{\psi}^{*} &= \sum_{h=0,1}\{\hat{Y}^{*}(0,1,h;\Gamma_{\psi}^{\mathrm{WIE}})-\hat{Y}^{*}(0,0,h;\Gamma_{\psi}^{\mathrm{WIE}})\},\\
\hat{\mathrm{BIE}}_{\psi}^{*} &= \sum_{s=0,1}\{\hat{Y}^{*}(0,s,1;\Gamma_{\psi}^{\mathrm{BIE}})-\hat{Y}^{*}(0,s,0;\Gamma_{\psi}^{\mathrm{BIE}})\}.
\end{aligned}$$

Choice of weights

- Choose weights to target policy ψ for different effects:

$$\gamma_{i,\psi}^{\mathrm{DE}}(a,s,h)=\pi_{i,\psi}(\cdot,s,h),\quad \gamma_{i,\psi}^{\mathrm{WIE}}(a,s,h)=\pi_{i,\psi}(\cdot,\cdot,h),\quad \gamma_{i,\psi}^{\mathrm{BIE}}(a,s,h)=\pi_{i,\psi}(\cdot,s,\cdot).$$

- All of the above weights can be simulated based on the design.
- Can be used to construct estimators for the policy-specific effects: for $* \in \{\mathrm{ht}, \mathrm{haj}\}$,

$$\begin{aligned}
\hat{\mathrm{DE}}_{\psi}^{*} &= \sum_{s,h=0,1}\{\hat{Y}^{*}(1,s,h;\Gamma_{\psi}^{\mathrm{DE}})-\hat{Y}^{*}(0,s,h;\Gamma_{\psi}^{\mathrm{DE}})\},\\
\hat{\mathrm{WIE}}_{\psi}^{*} &= \sum_{h=0,1}\{\hat{Y}^{*}(0,1,h;\Gamma_{\psi}^{\mathrm{WIE}})-\hat{Y}^{*}(0,0,h;\Gamma_{\psi}^{\mathrm{WIE}})\},\\
\hat{\mathrm{BIE}}_{\psi}^{*} &= \sum_{s=0,1}\{\hat{Y}^{*}(0,s,1;\Gamma_{\psi}^{\mathrm{BIE}})-\hat{Y}^{*}(0,s,0;\Gamma_{\psi}^{\mathrm{BIE}})\}.
\end{aligned}$$

- Marginal direct/indirect effects are a special case with the implemented policy.

Outline

Asymptotic variance: Horvitz–Thompson

- For the Horvitz–Thompson estimator:

$$
\begin{aligned}
\mathrm{var}\{\hat{Y}^{\mathrm{ht}}(a,s,h;\Gamma)\} = {} & \underbrace{n^{-2}\sum_{i=1}^{n}\frac{1-\pi_i(a,s,h)}{\pi_i(a,s,h)}\{\gamma_i(a,s,h)Y_i(a,s,h)\}^2}_{\text{contribution from the variation of unit } i} \\
& + n^{-2}\sum_{i=1}^{n}\sum_{j\neq i}\frac{\pi_{ij}(a,s,h;a,s,h)-\pi_i(a,s,h)\pi_j(a,s,h)}{\pi_i(a,s,h)\pi_j(a,s,h)} \\
& \underbrace{\qquad\qquad \times\gamma_i(a,s,h)\gamma_j(a,s,h)Y_i(a,s,h)Y_j(a,s,h).}_{\text{contribution from the covariance between units due to interference}}
\end{aligned}
$$

Asymptotic variance: Horvitz–Thompson

- For the Horvitz–Thompson estimator:

$$\mathrm{var}\{\hat{Y}^{\mathrm{ht}}(a,s,h;\Gamma)\} = \underbrace{n^{-2}\sum_{i=1}^{n}\frac{1-\pi_i(a,s,h)}{\pi_i(a,s,h)}\{\gamma_i(a,s,h)Y_i(a,s,h)\}^2}_{\text{contribution from the variation of unit } i}$$
$$+ \underbrace{n^{-2}\sum_{i=1}^{n}\sum_{j\neq i}\frac{\pi_{ij}(a,s,h;a,s,h)-\pi_i(a,s,h)\pi_j(a,s,h)}{\pi_i(a,s,h)\pi_j(a,s,h)} \times \gamma_i(a,s,h)\gamma_j(a,s,h)Y_i(a,s,h)Y_j(a,s,h).}_{\text{contribution from the covariance between units due to interference}}$$

- Can be written compactly as:

$$\mathrm{var}\{\hat{Y}^{\mathrm{ht}}(a,s,h;\Gamma)\} = n^{-2}\boldsymbol{Y}(a,s,h)^{\mathrm{T}}\boldsymbol{\Gamma}(a,s,h)\boldsymbol{\Omega}(a,s,h)\boldsymbol{\Gamma}(a,s,h)\boldsymbol{Y}(a,s,h).$$

Asymptotic variance: Hájek estimator

- For the Hájek estimator:

Asymptotic variance: Hájek estimator

- For the Hájek estimator:
  - Same form of the Horvitz–Thompson variance, with $\boldsymbol{Y}(a, s, h)$ replaced by the **centered** potential outcomes $\tilde{\boldsymbol{Y}}(a, s, h)$.

Asymptotic variance: Hájek estimator

- For the Hájek estimator:
  - Same form of the Horvitz–Thompson variance, with $\boldsymbol{Y}(a, s, h)$ replaced by the **centered** potential outcomes $\tilde{\boldsymbol{Y}}(a, s, h)$.
  - The centered potential outcomes are:

$$\tilde{Y}_i(a, s, h) = Y_i(a, s, h) - \frac{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h) Y_i(a, s, h)}{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h)}.$$

Asymptotic variance: Hájek estimator

- For the Hájek estimator:
  - Same form of the Horvitz–Thompson variance, with $\boldsymbol{Y}(a, s, h)$ replaced by the **centered** potential outcomes $\tilde{\boldsymbol{Y}}(a, s, h)$.
  - The centered potential outcomes are:

$$\tilde{Y}_i(a, s, h) = Y_i(a, s, h) - \frac{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h) Y_i(a, s, h)}{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h)}.$$

  - Compactly: $n^{-2} \tilde{\boldsymbol{Y}}(a, s, h)^{\mathrm{T}} \boldsymbol{\Gamma}(a, s, h) \boldsymbol{\Omega}(a, s, h) \boldsymbol{\Gamma}(a, s, h) \tilde{\boldsymbol{Y}}(a, s, h)$.

Asymptotic variance: Hájek estimator

- For the Hájek estimator:
  - Same form of the Horvitz–Thompson variance, with $\boldsymbol{Y}(a,s,h)$ replaced by the **centered** potential outcomes $\tilde{\boldsymbol{Y}}(a,s,h)$.
  - The centered potential outcomes are:

$$\tilde{Y}_i(a,s,h) = Y_i(a,s,h) - \frac{n^{-1}\sum_{i=1}^{n}\gamma_i(a,s,h)Y_i(a,s,h)}{n^{-1}\sum_{i=1}^{n}\gamma_i(a,s,h)}.$$

  - Compactly: $n^{-2}\tilde{\boldsymbol{Y}}(a,s,h)^{\mathrm{T}}\boldsymbol{\Gamma}(a,s,h)\boldsymbol{\Omega}(a,s,h)\boldsymbol{\Gamma}(a,s,h)\tilde{\boldsymbol{Y}}(a,s,h)$.
- Conditions for the consistency of the estimators (Aronow and Samii, 2017).

Asymptotic variance: Hájek estimator

- For the Hájek estimator:
  - Same form of the Horvitz–Thompson variance, with $\boldsymbol{Y}(a, s, h)$ replaced by the **centered** potential outcomes $\tilde{\boldsymbol{Y}}(a, s, h)$.
  - The centered potential outcomes are:

$$\tilde{Y}_i(a, s, h) = Y_i(a, s, h) - \frac{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h) Y_i(a, s, h)}{n^{-1} \sum_{i=1}^{n} \gamma_i(a, s, h)}.$$

  - Compactly: $n^{-2} \tilde{\boldsymbol{Y}}(a, s, h)^{\mathrm{T}} \boldsymbol{\Gamma}(a, s, h) \boldsymbol{\Omega}(a, s, h) \boldsymbol{\Gamma}(a, s, h) \tilde{\boldsymbol{Y}}(a, s, h)$.
- Conditions for the consistency of the estimators (Aronow and Samii, 2017).
- The centering property explains why empirically the Hájek estimator is usually more stable.

Covariance structure

- Horvitz–Thompson covariance for $(a, s, h) \neq (a', s', h')$:

Covariance structure

- Horvitz–Thompson covariance for $(a,s,h) \neq (a',s',h')$:

$$
\begin{aligned}
&\operatorname{cov}\{\hat{Y}^{\mathrm{ht}}(a,s,h;\Gamma),\, \hat{Y}^{\mathrm{ht}}(a',s',h';\Gamma)\} \\
&= \underbrace{-n^{-2}\sum_{i=1}^{n}\gamma_i(a,s,h)\gamma_i(a',s',h')Y_i(a,s,h)Y_i(a',s',h')}_{\text{Covariance of different potential outcomes for the same unit } i} \\
&\quad \underbrace{+\, n^{-2}\sum_{i=1}^{n}\sum_{j\neq i}\frac{\pi_{ij}(a,s,h;a',s',h') - \pi_i(a,s,h)\pi_j(a',s',h')}{\pi_i(a,s,h)\pi_j(a',s',h')} \times \gamma_i(a,s,h)\gamma_j(a',s',h')Y_i(a,s,h)Y_j(a',s',h')}_{\text{Covariance of different potential outcomes between units}}
\end{aligned}
$$

Covariance structure

- Horvitz–Thompson covariance for $(a, s, h) \neq (a', s', h')$:

$$
\begin{aligned}
&\operatorname{cov}\{\hat{Y}^{\mathrm{ht}}(a, s, h; \Gamma), \hat{Y}^{\mathrm{ht}}(a', s', h'; \Gamma)\} \\
&= \underbrace{-n^{-2} \sum_{i=1}^{n} \gamma_i(a, s, h)\gamma_i(a', s', h') Y_i(a, s, h) Y_i(a', s', h')}_{\text{Covariance of different potential outcomes for the same unit } i} \\
&\quad + n^{-2} \sum_{i=1}^{n} \sum_{j \neq i} \underbrace{\frac{\pi_{ij}(a, s, h; a', s', h') - \pi_i(a, s, h)\pi_j(a', s', h')}{\pi_i(a, s, h)\pi_j(a', s', h')} \times \gamma_i(a, s, h)\gamma_j(a', s', h') Y_i(a, s, h) Y_j(a', s', h')}_{\text{Covariance of different potential outcomes between units}}
\end{aligned}
$$

- Hájek covariance: Same form but with centered potential outcomes $\tilde{Y}_i(a, s, h)$ and $\tilde{Y}_j(a', s', h')$

CLT and confidence intervals

- Ongoing: to develop central limit theorems and confidence intervals for the proposed estimators.
- Central limit theorem: use Stein's method to capture correlation due to network interference and characterize asymptotic distributions.
- Confidence interval: relies on constructing valid variance estimators.

Outline

Regression-based analysis

- What if we analyze the experiments in a factorial regression flavor?

Regression-based analysis

- What if we analyze the experiments in a factorial regression flavor?
- **Partially saturated regression:**

$$Y_i \sim 1 + A_i + \tilde{S}_i + \tilde{H}_i + A_i\tilde{S}_i + A_i\tilde{H}_i,$$

where $\tilde{S}_i$ and $\tilde{H}_i$ are centered indicators using population means. Then:

Regression-based analysis

- What if we analyze the experiments in a factorial regression flavor?
- **Partially saturated regression:**

$$Y_i \sim 1 + A_i + \tilde{S}_i + \tilde{H}_i + A_i\tilde{S}_i + A_i\tilde{H}_i,$$

  where $\tilde{S}_i$ and $\tilde{H}_i$ are centered indicators using population means. Then:
  - Coefficient on A_i targets the marginal direct effect $\textsc{de}$.

Regression-based analysis

- What if we analyze the experiments in a factorial regression flavor?
- **Partially saturated regression:**

$$Y_i \sim 1 + A_i + \tilde{S}_i + \tilde{H}_i + A_i\tilde{S}_i + A_i\tilde{H}_i,$$

  where $\tilde{S}_i$ and $\tilde{H}_i$ are centered indicators using population means. Then:
  - Coefficient on A_i targets the marginal direct effect DE.
  - Centering yields "overlap" versions of indirect effects; exact WIE and BIE recovery requires weighting.

Regression-based analysis

- What if we analyze the experiments in a factorial regression flavor?
- **Partially saturated regression:**

$$Y_i \sim 1 + A_i + \tilde{S}_i + \tilde{H}_i + A_i\tilde{S}_i + A_i\tilde{H}_i,$$

  where $\tilde{S}_i$ and $\tilde{H}_i$ are centered indicators using population means. Then:
  - Coefficient on A_i targets the marginal direct effect DE.
  - Centering yields "overlap" versions of indirect effects; exact WIE and BIE recovery requires weighting.
- See Basse and Feller (2018); Zhao and Ding (2022).

Ongoing work

- Covariate adjustment:
  - How to use pre-treatment covariates to enhance the efficiency of the estimators?
    - Demographic features, characteristics of the graph, etc.

Ongoing work

- Covariate adjustment:
  - How to use pre-treatment covariates to enhance the efficiency of the estimators?
    - Demographic features, characteristics of the graph, etc.
  - Many classic discussions on *model-assisted* regression adjustment using pretreatment covariates, **without assuming the true outcome model**.

Ongoing work

- Covariate adjustment:
  - How to use pre-treatment covariates to enhance the efficiency of the estimators?
    - Demographic features, characteristics of the graph, etc.
  - Many classic discussions on *model-assisted* regression adjustment using pretreatment covariates, **without assuming the true outcome model**.
  - Adapting the framework to the current setting is a future step to enhance the analysis.

Ongoing work

- Covariate adjustment:
  - How to use pre-treatment covariates to enhance the efficiency of the estimators?
    - Demographic features, characteristics of the graph, etc.
  - Many classic discussions on *model-assisted* regression adjustment using pretreatment covariates, **without assuming the true outcome model**.
  - Adapting the framework to the current setting is a future step to enhance the analysis.
- Experimental validation:
  - The Cash Transfer experiment has a publicly available dataset.
  - Working on experimental validation with this study.

References I

Ali, M., Emch, M., von Seidlein, L., Yunus, M., Sack, D. A., Rao, M., Holmgren, J., and Clemens, J. D. (2005). Herd immunity conferred by killed oral cholera vaccines in bangladesh: a reanalysis. *The Lancet*, 366(9479):44–49.

Aronow, P. M. and Samii, C. (2017). Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947.

Baird, S., Bohren, J. A., McIntosh, C., and Özler, B. (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, 100(5):844–860.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015). The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7(1):22–53.

Barrera-Osorio, F., Linden, L. L., and Saavedra, J. E. (2019). Long-term impacts of alternative approaches to increase schooling: Evidence from a randomized experiment in colombia. *American Economic Journal: Applied Economics*, 11(3):33–54.

Basse, G. and Feller, A. (2018). Analyzing two-stage experiments in the presence of interference. *Journal of the American Statistical Association*, 113(521):41–55.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2):531–580.

References II

Duflo, E., Dupas, P., and Kremer, M. (2015). Education, hiv, and early fertility: Experimental evidence from kenya. *American Economic Review*, 105(9):2757–2797.

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. (2022). General equilibrium effects of cash transfers: experimental evidence from kenya. *Econometrica*, 90(6):2603–2643.

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

Miguel, E. and Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Pradhan, M., Suryadarma, D., Beatty, A., Wong, M., Alishjabana, A., Gaduh, A., and Artha, R. (2014). Improving educational quality through enhancing community participation: Results from a randomized field experiment in indonesia. *American Economic Journal: Applied Economics*, 6(2):105–126.

Zhao, A. and Ding, P. (2022). Regression-based causal inference with factorial experiments: estimands, model specifications and design-based properties. *Biometrika*, 109(3):799–815.