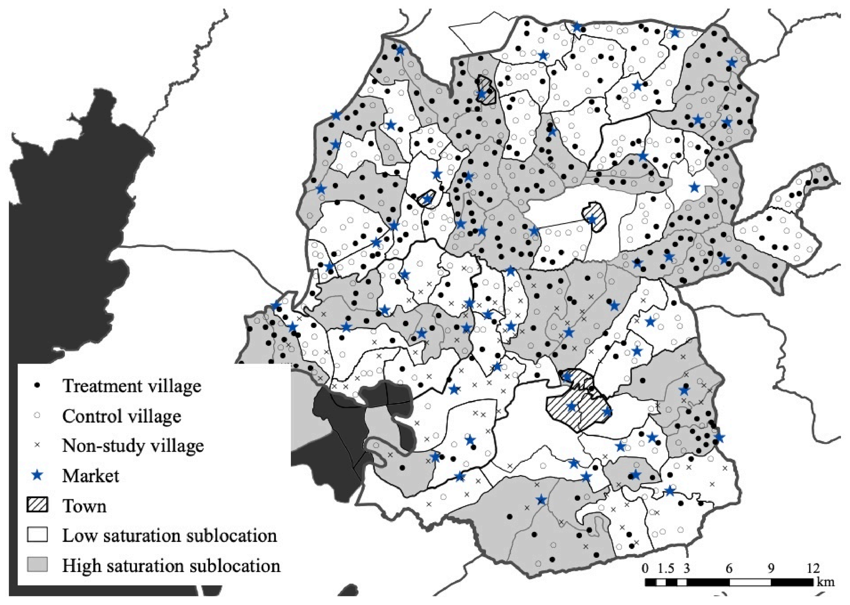
Estimating within-cluster and between-cluster spillover effects in randomized saturation designs

Sizhu Lu*, Lei Shi*, and Peng Ding

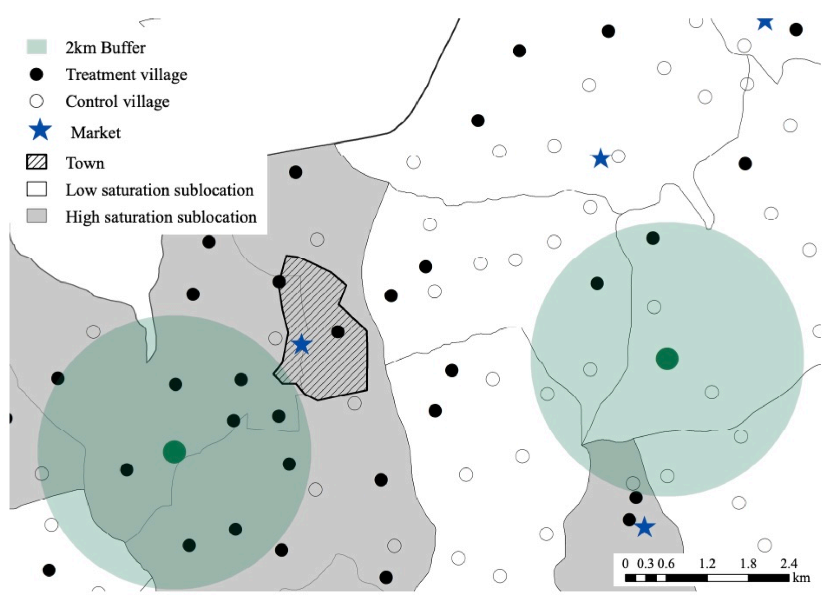
University of California, Berkeley

Motivating example



Large-scale cash transfer experiment in rural Kenya (Egger et al., 2022)

- Randomized saturation designs randomize **clusters to saturation levels** and then **units within clusters** to treatment.
- Common assumption in prior work: spillovers **only** occur within clusters, no interference between clusters.
- In practice, geography (markets, labor, mobility) induces interference *within* and *between* clusters.



Spatial exposure to treatment (Egger et al., 2022)

- We give a design-based framework for direct effects, within-cluster spillovers, and between-cluster spillovers under interference.

Randomized saturation design

- Two-stage *randomized saturation* design:
  - **Stage 1, cluster level:** Randomize clusters (communities, schools, villages, networks) to different saturation levels (high vs. low coverage).
  - **Stage 2, unit level:** Randomize individual units within each cluster to treatment, according to the assigned saturation rate.
- Finite population of n units (villages), indexed by i.
- $A_i \in \{0,1\}$: binary treatment indicator for village i; $\boldsymbol{A} = (A_1, \ldots, A_n)^{\mathrm{t}}$.
- $Y_i(\boldsymbol{a})$ for $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{t}}$: potential outcome of village i.
- Finite population regime: $\{Y_i(\boldsymbol{a})\}$ fixed; randomness comes from $\boldsymbol{A}$.
- Without further structure, 2^n potential outcomes per unit.

Exposure mapping and causal estimands

- Adopt an exposure mapping to reduce dimensionality:
$$d_i(\boldsymbol{A}) = (A_i, S_i, H_i),$$
where S_i summarizes the *within-sublocation* treatments and H_i summarizes the *between-sublocation* (geographically nearby) treatments.
- Running example used throughout (binary S_i, H_i):
  - 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 $Y_i(a, s, h)$.
- Eight potential outcomes per unit $\Rightarrow$ a 2^3 factorial structure with network-induced dependence.

Causal estimands: direct and indirect effects

Conditional effects

- Conditional direct effect of A_i, holding (S_i, H_i) fixed at (s, h):
$$\mathsf{DE}(s,h) = n^{-1}\sum_{i=1}^n Y_i(1,s,h) - n^{-1}\sum_{i=1}^n Y_i(0,s,h).$$
- Within-cluster conditional indirect effect of S_i, holding $H_i = h$:
$$\mathsf{WIE}(h) = n^{-1}\sum_{i=1}^n Y_i(0,1,h) - n^{-1}\sum_{i=1}^n Y_i(0,0,h).$$
- Between-cluster conditional indirect effect, holding $S_i = s$:
$$\mathsf{BIE}(s) = n^{-1}\sum_{i=1}^n Y_i(0,s,1) - n^{-1}\sum_{i=1}^n Y_i(0,s,0).$$

In-policy marginal effects: Aggregate the conditional effects by marginalizing over the exposure distribution induced by the implemented policy.

- Marginal direct effect of A_i:
$$\mathsf{DE} = n^{-1}\sum_{i=1}^n \underbrace{E\{Y_i(1,S_i,H_i) \mid A_i = 1\}}_{\text{marginalize over } (S_i,H_i) \text{ given } A_i=1} - n^{-1}\sum_{i=1}^n E\{Y_i(0,S_i,H_i) \mid A_i = 0\}.$$
- Within-cluster marginal indirect effect:
$$\mathsf{WIE} = n^{-1}\sum_{i=1}^n E\{Y_i(0,1,H_i) \mid (A_i,S_i) = (0,1)\} - n^{-1}\sum_{i=1}^n E\{Y_i(0,0,H_i) \mid (A_i,S_i) = (0,0)\}.$$
- Between-cluster marginal indirect effect:
$$\mathsf{BIE} = n^{-1}\sum_{i=1}^n E\{Y_i(0,S_i,1) \mid (A_i,H_i) = (0,1)\} - n^{-1}\sum_{i=1}^n E\{Y_i(0,S_i,0) \mid (A_i,H_i) = (0,0)\}.$$

Policy-specific effects

- For a given treatment policy ψ, define
$$\mathsf{DE}_\psi = n^{-1}\sum_{i=1}^n E_\psi\{Y_i(1,S_i,H_i) \mid A_i = 1\} - n^{-1}\sum_{i=1}^n E_\psi\{Y_i(0,S_i,H_i) \mid A_i = 0\},$$
  - similar marginalizations for WIE_ψ and BIE_ψ.
  - subscript ψ: $E_\psi(\cdot \mid \cdot)$ is taken under the conditional distribution induced by ψ.
- Comparison between different policies ψ_1 and ψ_2, e.g., $\mathsf{DE}_{\psi_1} - \mathsf{DE}_{\psi_2}$.

Estimation: inverse propensity score weighting

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

- Define $\mathbb{I}_i(a,s,h) = 1\{A_i = a, S_i = s, H_i = h\}$ and $\pi_i(a,s,h) = \mathrm{pr}(A_i = a, S_i = s, H_i = h)$.
- Horvitz–Thompson estimator:
$$\widehat{Y}^{\mathsf{ht}}(a,s,h) = n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i.$$
- Hájek estimator:
$$\widehat{Y}^{\mathsf{haj}}(a,s,h) = n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)} Y_i \Big/ n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)}{\pi_i(a,s,h)}.$$

Conditional effects: differences of $\widehat{Y}^*(a,s,h)$.

- e.g., $\widehat{\mathsf{DE}}^*(s,h) = \widehat{Y}^*(1,s,h) - \widehat{Y}^*(0,s,h)$ for $* \in \{\mathsf{ht}, \mathsf{haj}\}$.

Marginal policy-specific effects: reweighted estimators.

Reweighted estimators for policy-specific effects

- Different policies ψ induce different distributions of (A_i, S_i, H_i). To target ψ, we **reweight** the IPW estimators.
- $\Gamma = \{\gamma_i(a,s,h) : i = 1, \ldots, n,\ a, s, h \in \{0,1\}\}$: a set of prespecified unit-level weights.
- Horvitz–Thompson estimators:
$$\widehat{Y}^{\mathsf{ht}}(a,s,h;\Gamma) = n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)\,\gamma_i(a,s,h)}{\pi_i(a,s,h)} Y_i.$$
- Hájek estimators:
$$\widehat{Y}^{\mathsf{haj}}(a,s,h;\Gamma) = n^{-1}\sum_{i=1}^n \gamma_i(a,s,h) \cdot n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)\gamma_i(a,s,h)}{\pi_i(a,s,h)} Y_i \Big/ n^{-1}\sum_{i=1}^n \frac{\mathbb{I}_i(a,s,h)\gamma_i(a,s,h)}{\pi_i(a,s,h)}.$$
- Choice of weights: policy-induced conditional probabilities as weights
$$\gamma_{i,\psi}^{\mathsf{DE}}(a,s,h) = \mathrm{pr}_\psi(S_i = s, H_i = h \mid A_i = a),$$
similar for $\gamma_{i,\psi}^{\mathsf{WIE}}(a,s,h)$ and $\gamma_{i,\psi}^{\mathsf{BIE}}(a,s,h)$.

Asymptotic theory and inference

Regularity conditions

- Bounded potential outcomes and positivity of all relevant exposure probabilities.
- Bounded cluster degree: each unit has a finite local dependence neighborhood.
- Bounded order of dependence: exposures become independent beyond a fixed graph distance.

Closed-form variance: Horvitz–Thompson cell mean

Let $\boldsymbol{Y}(a,s,h)$ collect the finite-population potential outcomes, $\boldsymbol{\Gamma}$ be diagonal with entries $\gamma_i(a,s,h)$, and $\boldsymbol{\Lambda}$ encode first- and second-order inclusion covariance. Then
$$\mathrm{var}\{\widehat{Y}^{\mathsf{ht}}(a,s,h;\Gamma)\} = n^{-2}\boldsymbol{Y}(a,s,h)^\top \boldsymbol{\Gamma}(a,s,h)\boldsymbol{\Lambda}(a,s,h)\boldsymbol{\Gamma}(a,s,h)\boldsymbol{Y}(a,s,h).$$
For Hájek estimators, use the same form with **centered** potential outcomes.

Joint consistency and central limit theorem

Under the regularity conditions and non-degeneracy of the asymptotic variance, for $* \in \{\mathsf{ht}, \mathsf{haj}\}$:

(a) Consistency: $\widehat{\boldsymbol{Y}}^*_\Gamma - \bar{\boldsymbol{Y}}_\Gamma \xrightarrow{p} 0$.

(b) Joint asymptotic normality:
$$\mathsf{acov}(\widehat{\boldsymbol{Y}}^*_\Gamma)^{-1/2}(\widehat{\boldsymbol{Y}}^*_\Gamma - \bar{\boldsymbol{Y}}_\Gamma) \xrightarrow{d} \mathcal{N}(0, I_8),$$
where $\widehat{\boldsymbol{Y}}^*_\Gamma$ is the joint vector across all 8 exposure cells.

- All proposed estimators (conditional, in-policy marginal, policy-specific) are consistent and asymptotically normal by the continuous mapping theorem and the delta method.

Conservative variance estimators

- Define $\boldsymbol{\Omega}(a,s,h)$:
$$\boldsymbol{\Omega}_{ij}(a,s,h) = \frac{\pi_{ij}(a,s,h;a,s,h) - \pi_i(a,s,h)\pi_j(a,s,h)}{\pi_{ij}(a,s,h;a,s,h)},$$
  - $\pi_{ii}(a,s,h;a,s,h) = \pi_i(a,s,h)$, so diagonal $= 1 - \pi_i(a,s,h)$.
- Single-cell variance estimator (consistent for $\mathrm{avar}\{\widehat{Y}^*(a,s,h;\Gamma)\}$):
$$\widehat{\mathsf{se}}^2\{\widehat{Y}^*(a,s,h;\Gamma)\} = n^{-2}(\widehat{\boldsymbol{Y}}^*)^\top \boldsymbol{\Gamma}(a,s,h)\boldsymbol{\Omega}(a,s,h)\boldsymbol{\Gamma}(a,s,h)\widehat{\boldsymbol{Y}}^*,$$
  - $\widehat{Y}_i^{\mathsf{ht}} = \mathbb{I}_i(a,s,h)Y_i/\pi_i(a,s,h)$; $\widehat{Y}_i^{\mathsf{haj}}$ centered version.
- Cross-level covariance is unidentifiable for *contrasts*: Cauchy–Schwarz to obtain a **conservative** bound.

Marginal effects can mask heterogeneity

Conditional direct and between-cluster indirect effects

Direct effects $\widehat{\mathsf{DE}}^*(s,h)$

(s,h)	Hájek Profit	Hájek Revenue	Covariate-adjusted Profit	Covariate-adjusted Revenue
$(0,0)$	-360	-46	-101	217
	(909)	(1,143)	(883)	(1,117)
$(0,1)$	-558	300	-835	-124
	(616)	(877)	(559)	(696)
$(1,0)$	**3,539*****	**4,159*****	**2,817****	**3,043*****
	(1,263)	(1,247)	(1,268)	(1,167)
$(1,1)$	-227	140	37	481
	(688)	(825)	(677)	(776)

Between-cluster indirect effects $\widehat{\mathsf{BIE}}^*(a,s)$

(a,s)	Hájek Profit	Hájek Revenue	Covariate-adjusted Profit	Covariate-adjusted Revenue
$(0,0)$	**1,330****	**1,950*****	**1,442*****	**1,998*****
	(554)	(645)	(511)	(555)
$(0,1)$	375	674	42	-92
	(954)	(1,128)	(967)	(1,138)
$(1,0)$	1,132	2,296*	708	1,658
	(971)	(1,374)	(930)	(1,258)
$(1,1)$	**−3,391*****	**−3,344*****	**−2,739*****	**−2,654*****
	(997)	(944)	(978)	(805)

- Control villages in low-saturation sublocations benefit from proximity to treated villages, possibly due to increased economic activity.
- Treated villages in high-saturation sublocations are negatively affected by additional geographic exposure: negative competitive spillover.

In-policy marginal effects

	Hájek Profit	Hájek Revenue	Covariate-adjusted Profit	Covariate-adjusted Revenue
$\widehat{\mathsf{DE}}$	336	811	289	656
	(1,052)	(1,242)	(995)	(1,119)
$\widehat{\mathsf{BIE}}$	358	601	436	565
	(1,138)	(1,321)	(1,060)	(1,160)

- No marginal effect significant: averaging *dilutes* the highly heterogeneous conditional signals.