Inference for Two-stage Experiments under Covariate-Adaptive Randomization

Jizhou Liu, Chicago Booth & Amazon Science

Randomized controlled trials (RCTs) ubiquitous throughout sciences

Interference widely encountered in experimental designs

- Medication: Miguel and Kremer (2004)
- Information treatment: Duflo and Saez (2003)
- Cash transfer: Haushofer and Shapiro (2016)

Key challenges related to interference:

- Naive estimators bias primary effects:
  - Example: mean difference of vaccinated vs unvaccinated
  - Primary effects: total impact of vaccination (defined later)
- Emergence of additional parameters of interest, such as spillover effects

Two-stage randomized experiments:

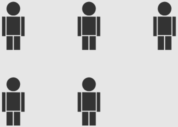
- Common approach to address interference
- Stage 1: Assigns treatment intensity at cluster level
- Stage 2: Assigns individual-level treatment within clusters

Example: Duflo and Saez (2003)

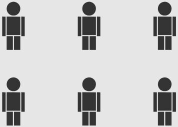
- Stage 1: Assign 50% or 0% treatment intensity to departments
- Stage 2: Assign $20 to treated individuals in "50% departments"

Example from Duflo and Saez (2003)

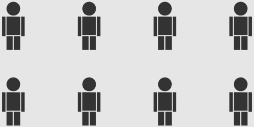
Department of Physics



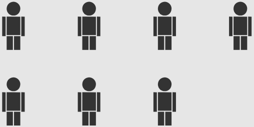
Department of Mathematics



Department of Economics

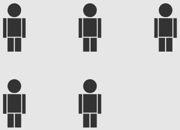


Department of Statistics

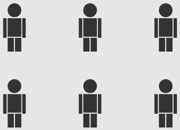


Example from Duflo and Saez (2003)

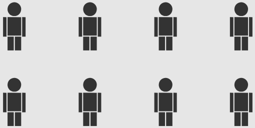
Department of Physics



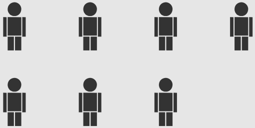
Department of Mathematics



Department of Economics

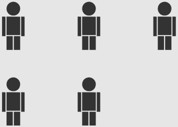


Department of Statistics

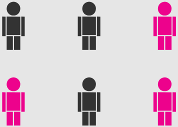


Example from Duflo and Saez (2003)

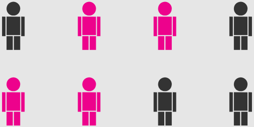
Department of Physics



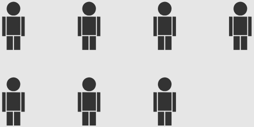
Department of Mathematics



Department of Economics



Department of Statistics



Covariate-adaptive randomization especially common in two-stage experiments:

Stratification: Hidrobo et al. (2016); Banerjee et al. (2021); Rogers and Feller (2018); Foos and de Rooij (2017); Muralidharan and Sundararaman (2015)

Matching: Duflo and Saez (2003); Beuermann et al. (2015); Ichino and Schündeln (2012); Kinnan et al. (2020); Malani et al. (2021)

Why use covariates? To improve efficiency

Applicable to both stages

Two-stage with matched pairs:

- Pair clusters
- Assign 50% treatment to one cluster, 0% to the other
- Pair units within the 50% cluster
- Randomly assign one unit to treatment within each pair

Introduction (cont.)

Goals for today:

- Assess validity of common inference methods from the literature
- Propose asymptotically exact inference methods
- Optimize covariate use in two-stage designs
- Introduce ex post covariate adjustment

Two-stage experiments (ignoring covariates)

Basse and Feller (2018); Basse et al. (2019); Imai et al. (2021)

Partial population experiments

Tortarolo et al. (2023); Baird et al. (2018); Hirano and Hahn (2010)

Split-plot design

Zhao and Ding (2022); Shi et al. (2022)

Cluster randomized trials

Bugni et al. (2022); Bai et al. (2022)

Index clusters (e.g. departments) by $g = 1, \ldots, G$ and units in gth cluster by $i = 1, \ldots, N_g$. Some notation:

$$
\begin{aligned}
Y_{i,g} &= \text{observed outcome of unit } i \text{ in cluster } g \text{ (e.g. participation)} \\
Z_{i,g} &= \text{treatment assignment of unit } i \text{ in cluster } g \text{ (e.g. \$20 check)} \\
H_g &= \text{fraction of treated units in cluster } g \\
N_g &= \text{size of cluster } g
\end{aligned}
$$

Setup and Notation (cont.)

Binary treatment:

Intensity levels: $H_g \in \{0, \pi_2\}$, Treatment status: $Z_{i,g} \in \{0, 1\}$

Stage 1:

π_1 fraction of clusters with $H_g = \pi_2$, remainder with $H_g = 0$

Stage 2:

Within each treated cluster, π_2 fraction of units with $Z_{i,g} = 1$, others with $Z_{i,g} = 0$

Potential outcomes:

$$Y_{i,g}(z,h) \in \{Y_{i,g}(1,\pi_2), Y_{i,g}(0,\pi_2), Y_{i,g}(0,0)\}$$

where

z: Unit-level treatment status, $z \in \{0,1\}$.

h: Fraction of treated units in the cluster, $h \in \{0,\pi_2\}$.

Observed outcome: $Y_{i,g} = Y_{i,g}(Z_{i,g}, H_g)$.

Equally-weighted cluster-level average treatment effect:

$$\gamma^P := E\left[\frac{1}{N_g}\sum_{i=1}^{N_g}(Y_{i,g}(1,\pi_2) - Y_{i,g}(0,0))\right] \quad \text{(primary effect)}$$

$$\gamma^S := E\left[\frac{1}{N_g}\sum_{i=1}^{N_g}(Y_{i,g}(0,\pi_2) - Y_{i,g}(0,0))\right] \quad \text{(spillover effect)}$$

$\Longrightarrow$ *clusters are of interest.*

Size-weighted cluster-level average treatment effect:

$$\theta^P := E\left[\frac{1}{E[N_g]}\sum_{i=1}^{N_g}(Y_{i,g}(1,\pi_2) - Y_{i,g}(0,0))\right] \quad \text{(primary effect)}$$

$$\theta^S := E\left[\frac{1}{E[N_g]}\sum_{i=1}^{N_g}(Y_{i,g}(0,\pi_2) - Y_{i,g}(0,0))\right] \quad \text{(spillover effect)}$$

$\Longrightarrow$ *units in the clusters are of interest.*

Weighted primary and spillover effects:

Definition

Primary effect:

$$\theta^P_\omega := E\left[\omega_g\left(\frac{1}{N_g}\sum_{i=1}^{N_g} Y_{i,g}(1,\pi_2) - Y_{i,g}(0,0)\right)\right]$$

Spillover effect:

$$\theta^S_\omega := E\left[\omega_g\left(\frac{1}{N_g}\sum_{i=1}^{N_g} Y_{i,g}(0,\pi_2) - Y_{i,g}(0,0)\right)\right]$$

Equally-weighted: $\omega_g = 1$; Size-weighted: $\omega_g = N_g/E[N_g]$.

OLS based on the model:

$$Y_{i,g} = \alpha + \beta_1 Z_{i,g} + \beta_2 L_{i,g} + \epsilon_{i,g} \ ,$$

where $L_{i,g} = I\{H_g = \pi_2\}(1 - Z_{i,g})$.

1. Standard OLS yields size-weighted estimators $\hat{\theta}^P$ and $\hat{\theta}^S$:

$$\hat{\theta}^P \xrightarrow{p} \theta^P, \quad \hat{\theta}^S \xrightarrow{p} \theta^S \ .$$

2. OLS weighted by $1/N_g$ yields equally-weighted estimators $\tilde{\theta}^P$ and $\tilde{\theta}^S$:

$$\tilde{\theta}^P \xrightarrow{p} \gamma^P, \quad \tilde{\theta}^S \xrightarrow{p} \gamma^S \ .$$

Example:

$$\hat{\theta}^P = \frac{1}{N_1} \sum_{1 \le g \le G} I\{H_g = \pi_2\} N_g \bar{Y}_g^1 - \frac{1}{N_0} \sum_{1 \le g \le G} I\{H_g = 0\} N_g \bar{Y}_g^1$$

where

$N_1 = \#$ units in treated clusters

$N_0 = \#$ units in control clusters

$$\bar{Y}_g^1 = \begin{cases} \text{average outcome of treated units,} & \text{if cluster } g \text{ is treated} \\ \text{average outcome of control units,} & \text{if cluster } g \text{ is control} \end{cases}$$

Caveats for regression-based estimators:

- Is everyone observed in each cluster?
  - OLS inconsistent for θ^P and θ^S if only a subset of units observed
  - Use weighted OLS: $\frac{\text{cluster size}}{\text{sample size}}$ to recover θ^P and θ^S
- Strata fixed effects:
  - Avoid adding strata fixed effects in the regression

Preview:

- No covariates used in design:
  - "OLS + Cluster robust standard errors" is asymptotically exact
- Covariates used for stratification or matching:
  - Requires further asymptotic analysis
  - Notation: C_g (cluster-level), $X_{i,g}$ (individual-level)

Stage 1:

n strata, each of size k (clusters grouped by $S_g = S(C_g, N_g)$).

Denote the j-th stratum by λ_j; example:

$$\{1, \ldots, G\} \xrightarrow{\text{sort by } S_g} \{\underbrace{10, 2, 5}_{\lambda_1}, \underbrace{21, 99, 6}_{\lambda_2}, \ldots, \underbrace{47, 23, 11}_{\lambda_n}\}$$

Stage 2: similar to Stage 1.

Assumption 1

Strata satisfy:

$$\frac{1}{n} \sum_{1 \le j \le n} \max_{i,k \in \lambda_j} \|S_i - S_k\|^2 \xrightarrow{P} 0 \ .$$

Theorem 1

Under previous assumptions + regularity + sampling procedure conditions, as $G \to \infty$,

$$\sqrt{G}\left(\hat{\theta}^P - \theta^P\right) \xrightarrow{d} \mathcal{N}(0, V(1)) \ ,$$
$$\sqrt{G}\left(\hat{\theta}^S - \theta^S\right) \xrightarrow{d} \mathcal{N}(0, V(0)) \ ,$$

where

$$V(z) = \frac{1}{\pi_1}\operatorname{Var}[\tilde{Y}_g(z,\pi_2)] + \frac{1}{1-\pi_1}\operatorname{Var}[\tilde{Y}_g(0,0)] - \pi_1(1-\pi_1)E\left[\left(\frac{1}{\pi_1}E[\tilde{Y}_g(z,\pi_2) \mid S_g] + \frac{1}{1-\pi_1}E[\tilde{Y}_g(0,0) \mid S_g]\right)^2\right] \ ,$$

with $\tilde{Y}_g(z,\pi_2)$ as the demeaned, size-weighted average outcomes.

$$V(z) = \underbrace{\frac{1}{\pi_1}\,\mathrm{Var}[\tilde{Y}_g(z,\pi_2)] + \frac{1}{1-\pi_1}\,\mathrm{Var}[\tilde{Y}_g(0,0)]}_{:=V_1(z)}$$
$$\underbrace{-\,\pi_1(1-\pi_1)E\left[\left(\frac{1}{\pi_1}E[\tilde{Y}_g(z,\pi_2)\mid S_g] + \frac{1}{1-\pi_1}E[\tilde{Y}_g(0,0)\mid S_g]\right)^2\right]}_{:=V_2(z)}$$

Cluster-robust variance estimator:

- Generally conservative: $\hat{V}_{CR}(z) \xrightarrow{p} V_1(z) \geq V(z)$.
- Consistent if design is not covariate-adaptive, or S_g is irrelevant.

$$V(z) = \underbrace{\frac{1}{\pi_1}\,\mathrm{Var}[\tilde{Y}_g(z,\pi_2)] + \frac{1}{1-\pi_1}\,\mathrm{Var}[\tilde{Y}_g(0,0)]}_{:=V_1(z)}$$

$$- \underbrace{\pi_1(1-\pi_1)E\left[\left(\frac{1}{\pi_1}E[\tilde{Y}_g(z,\pi_2) \mid S_g] + \frac{1}{1-\pi_1}E[\tilde{Y}_g(0,0) \mid S_g]\right)^2\right]}_{:=V_2(z)}$$

Efficiency gain from the first-stage stratification/matching

$$V(z) = \underbrace{\frac{1}{\pi_1} \mathrm{Var}[\tilde{Y}_g(z, \pi_2)] + \frac{1}{1-\pi_1} \mathrm{Var}[\tilde{Y}_g(0,0)]}_{:=V_1(z)}$$

$$\underbrace{- \pi_1(1-\pi_1)E\left[\left(\frac{1}{\pi_1}E[\tilde{Y}_g(z,\pi_2) \mid S_g] + \frac{1}{1-\pi_1}E[\tilde{Y}_g(0,0) \mid S_g]\right)^2\right]}_{:=V_2(z)}$$

Efficiency gain from second-stage stratification/matching

$\tilde{Y}_g(1, \pi_2)$: average of treated units (depends on $Z_{i,g}$)

Estimation and Inference (cont.)

$$V(z) = \underbrace{\frac{1}{\pi_1}\operatorname{Var}[\tilde{Y}_g(z,\pi_2)] + \frac{1}{1-\pi_1}\operatorname{Var}[\tilde{Y}_g(0,0)]}_{V_1(z)}$$
$$\underbrace{- \pi_1(1-\pi_1)E\left[\left(\frac{1}{\pi_1}E[\tilde{Y}_g(z,\pi_2) \mid S_g] + \frac{1}{1-\pi_1}E[\tilde{Y}_g(0,0) \mid S_g]\right)^2\right]}_{V_2(z)}$$

Relative importance:

First stage: $O(1)$ vs. Second stage: $O(1/N)$

N: smallest cluster size

How to estimate $V(z)$ consistently?

First, define

$$\tilde{Y}_g^z = \frac{N_g}{\frac{1}{G}\sum_{1 \le g \le G} N_g}\left(\bar{Y}_g^z - \frac{\frac{1}{G_g}\sum_{1 \le j \le G} \bar{Y}_j^z I\{H_g = H_j\} N_j}{\frac{1}{G}\sum_{1 \le j \le G} N_j}\right) ,$$

$\tilde{Y}_g^z$ acts like a “feasible” version of $\tilde{Y}_g(z,h)$:

$$H_g = h \Longrightarrow \tilde{Y}_g^z \approx \tilde{Y}_g(z,h).$$

Using this, “easy” to estimate quantities like $E[\tilde{Y}_g^2(z,h)]$ using, e.g.,

$$\frac{1}{G\pi_1}\sum_{1 \le g \le G} (\tilde{Y}_g^z)^2 I\{H_g = h\} .$$

"Harder" to estimate quantities like $E\left[E[\tilde{Y}_g(z,h) \mid S_g]^2\right]$

Problem: Need two independent copies of $\tilde{Y}_g(z,h)$ within a stratum.

matched pair: observe one $\tilde{Y}_g(0,0)$ and one $\tilde{Y}_g(z,\pi_2)$.

Solution: Look at nearby pairs.

Strategy works if we assume adjacent pairs are suitably close:

Assumption 2

Strata satisfy:

$$\frac{1}{n}\sum_{1\le j\le \lfloor n/2\rfloor} \max_{i\in\lambda_{2j-1},k\in\lambda_{2j}} \|S_i - S_k\|^2 \xrightarrow{P} 0 \ .$$

For matched-pair design, this strategy leads to following estimator:

$$\hat{V}(z) = \hat{\tau}_G^2 - \frac{1}{2}\hat{\lambda}_G^2$$

where

$$\hat{\tau}_G^2 = \frac{2}{G}\sum_{j \leq G/2}\left(\sum_{i \in \lambda_j}\tilde{Y}_i^z I\{H_i = \pi_2\} - \sum_{i \in \lambda_j}\tilde{Y}_i^z I\{H_i = 0\}\right)^2$$

$$\hat{\lambda}_G^2 = \frac{4}{G}\sum_{j \leq G/4}\left(\sum_{i \in \lambda_{2j-1}}\tilde{Y}_i^z I\{H_i = \pi_2\} - \sum_{i \in \lambda_{2j-1}}\tilde{Y}_i^z I\{H_i = 0\}\right) \times \left(\sum_{i \in \lambda_{2j}}\tilde{Y}_i^z I\{H_i = \pi_2\} - \sum_{i \in \lambda_{2j}}\tilde{Y}_i^z I\{H_i = 0\}\right)$$

Summary:

Estimation:

OLS from the model:

$$Y_{i,g} = \alpha + \beta_1 Z_{i,g} + \beta_2 L_{i,g} + \epsilon_{i,g},$$

weighted by ω_g / N_g for θ^P_ω, θ^S_ω.

Inference: Adjusted t-test using variance estimator $\hat{V}(z)$.

Outline

Optimal matching via index function in the first-stage:

Theorem 2

$V(z)$ *is minimized when* $S_g = E\left[\frac{\tilde{Y}_g(z,\pi_2)}{\pi_1} + \frac{\tilde{Y}_g(0,0)}{1-\pi_1} \mid C_g, N_g\right]$.

Design Recommendations:

1. Match on N_g (predictive of scale $\tilde{Y}_g(z,h)$).
2. Match on sums of individual covariates, not averages (e.g., Duflo and Saez (2003), match on the number of participants, not participation rate).
3. Use estimated index function (with large pilot).

Under homogeneous covariance:

Theorem 3

$V(z)$ *is minimized by stratifying/matching on* $E\left[Y_{i,g}(z,\pi_2) \mid (X_{i,g} : 1 \leq i \leq N_g)\right]$.

Design Recommendation:

For highly correlated outcomes within clusters, match on both individual and neighbors' covariates.

Match/stratify on a small set of covariates; adjust for additional covariates ex-post.

Linear adjustment improves efficiency:

> Adjusted outcomes: $Y_{i,g} - (\psi_g - \bar{\psi}_G)'\hat{\beta}^P / N_g$ (for primary effects), where ψ_g is the sum of covariates in the g-th cluster.
>
> Estimate $\hat{\beta}^P$:
>
> 1. Compute pairwise differences for $\tilde{Y}_g^1 \bar{N}_G$ and ψ_g in each stratum.
> 2. Run linear regression on the pairwise differences.

Taking the covariate-adjusted estimator $\hat{\theta}^{P,adj}$ as an example:

$$\sqrt{G}\left(\hat{\theta}^{P,adj} - \theta^P\right) \xrightarrow{d} \mathcal{N}(0, V(1) - \kappa^2)\ ,$$

Efficiency improvement with

$$E\left[\mathrm{Cov}\left[\frac{1}{\pi_1}\tilde{Y}_g(1,\pi_2) + \frac{1}{1-\pi_1}\tilde{Y}_g(0,0), \psi_g \mid S_g\right]\right] \neq 0\ .$$

First-stage	Parameter	**Complete Rand.**	**Stratification**	**Matched-pair**
Complete Rand.	θ^P	1.0000	0.9803	0.9560
	θ^S	1.0000	0.9921	0.9596
Stratification	θ^P	0.8227	0.7601	0.7672
	θ^S	0.8244	0.7790	0.7693
Matched-pair	θ^P	0.0921	0.0873	0.0755
	θ^S	0.0997	0.0930	0.0757

Table: Ratio of MSE under various designs

Simulation (cont.)

First-stage	Parameter	**Complete Rand.**	**Stratification**	**Matched-pair**
Complete Rand.	θ^P	1.0000	0.9803	0.9560
	θ^S	1.0000	0.9921	0.9596
Stratification	θ^P	0.8227	0.7601	0.7672
	θ^S	0.8244	0.7790	0.7693
Matched-pair	θ^P	0.0921	0.0873	0.0755
	θ^S	0.0997	0.0930	0.0757

Table: Ratio of MSE under various designs

Simulation (cont.)

First-stage	Parameter	**Complete Rand.**	**Stratification**	**Matched-pair**
Complete Rand.	θ^P	1.0000	0.9803	0.9560
	θ^S	1.0000	0.9921	0.9596
Stratification	θ^P	0.8227	0.7601	0.7672
	θ^S	0.8244	0.7790	0.7693
Matched-pair	θ^P	0.0921	0.0873	0.0755
	θ^S	0.0997	0.0930	0.0757

Table: Ratio of MSE under various designs

Inference Method	Effect	**S** **C**	**S** **S**	**S** **MP**	**MP** **C**	**MP** **S**	**MP** **MP**
Adjusted t-test	Primary	0.048	0.045	0.058	0.052	0.057	0.051
	Spillover	0.049	0.046	0.061	0.047	0.058	0.050
OLS robust (standard t-test)	Primary	0.184	0.194	0.156	0.062	0.086	0.049
	Spillover	0.184	0.167	0.159	0.077	0.048	0.048
OLS cluster (clustered t-test)	Primary	0.000	0.000	0.000	0.000	0.000	0.000
	Spillover	0.000	0.000	0.000	0.000	0.000	0.000
OLS with strata fixed effects (robust)	Primary	0.209	0.196	0.179	0.100	0.106	0.077
	Spillover	0.201	0.184	0.177	0.113	0.100	0.075
OLS with strata fixed effects (clustered)	Primary	0.028	0.027	0.029	0.068	0.085	0.071
	Spillover	0.036	0.027	0.026	0.064	0.062	0.069

Table: Rejection probabilities under $\beta_1 = \beta_2 = 0$

Simulation (cont.)

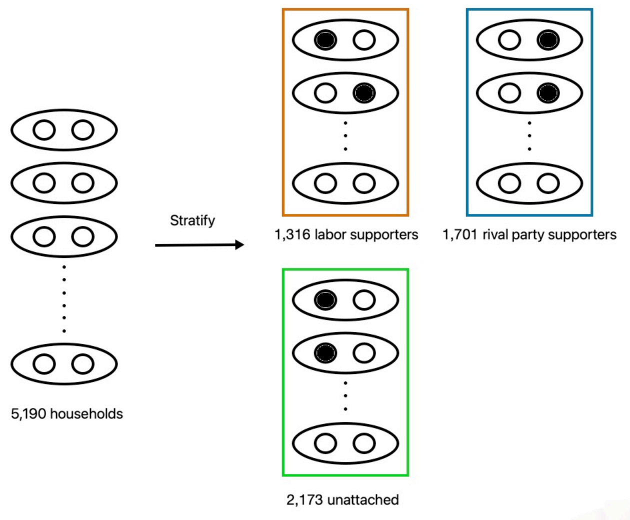
Inference Method	Effect	S C	S S	S MP	MP C	MP S	MP MP
Adjusted t-test	Primary	0.048	0.045	0.058	0.052	0.057	0.051
	Spillover	0.049	0.046	0.061	0.047	0.058	0.050
OLS robust (standard t-test)	Primary	0.184	0.194	0.156	0.062	0.086	0.049
	Spillover	0.184	0.167	0.159	0.077	0.048	0.048
OLS cluster (clustered t-test)	Primary	0.000	0.000	0.000	0.000	0.000	0.000
	Spillover	0.000	0.000	0.000	0.000	0.000	0.000
OLS with strata fixed effects (robust)	Primary	0.209	0.196	0.179	0.100	0.106	0.077
	Spillover	0.201	0.184	0.177	0.113	0.100	0.075
OLS with strata fixed effects (clustered)	Primary	0.028	0.027	0.029	0.068	0.085	0.071
	Spillover	0.036	0.027	0.026	0.064	0.062	0.069

Table: Rejection probabilities under $\beta_1 = \beta_2 = 0$

Foos and de Rooij (2017): randomized experiments in UK on social influence within households

- Two-stage randomization: households, then individuals
- Treatment: telephone encouragement to vote (PCC election)
- Outcome: turnout rates

Empirical Application (cont.)



	adjusted t-test	OLS robust	OLS cluster	OLS fe robust	OLS fe cluster
Primary	3.0488	3.0488	3.0488	2.9971	2.9971
	$\pm$ 2.2149	$\pm$ 2.0526	$\pm$ 2.2386	$\pm$ 2.0337	$\pm$ 2.2159
Spillover	4.5930	4.5930	4.5930	4.5413	4.5413
	$\pm$ 2.2500	$\pm$ 2.0885	$\pm$ 2.2715	$\pm$ 2.0719	$\pm$ 2.2509

Table: Point estimates and confidence intervals

OLS robust: narrower intervals (invalid).

OLS cluster: wider intervals (conservative).

OLS fe: different point estimates.

Two-stage experiments are common

Often involve covariate-adaptive randomization

Developed methods for inference for matching/stratification

Provided guidance for efficient experimental designs

Proposed covariate-adjusted estimator

References I

Bai, Y., Liu, J., Shaikh, A. M. and Tabord-Meehan, M. (2022). Inference in cluster randomized trials with matched pairs. URL `https://arxiv.org/abs/2211.14903`.

Baird, S., Bohren, J. A., McIntosh, C. and Özler, B. (2018). Optimal Design of Experiments in the Presence of Interference. *The Review of Economics and Statistics*, **100** 844–860.

Banerjee, A., Chattopadhyay, R., Duflo, E., Keniston, D. and Singh, N. (2021). Improving police performance in rajasthan, india: Experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy*, **13** 36–66.

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

Basse, G. W., Feller, A. and Toulis, P. (2019). Randomization tests of causal effects under interference. *Biometrika*, **106** 487–494.

Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O. and Cruz-Aguayo, Y. (2015). One laptop per child at home: Short-term impacts from a randomized experiment in peru. *American Economic Journal: Applied Economics*, **7** 53–80.

References II

Bugni, F., Canay, I., Shaikh, A. and Tabord-Meehan, M. (2022). Inference for cluster randomized experiments with non-ignorable cluster sizes. URL `https://arxiv.org/abs/2204.08356`.

Duflo, E. and Saez, E. (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment*. *The Quarterly Journal of Economics*, **118** 815–842.

Foos, F. and de Rooij, E. A. (2017). All in the family: Partisan disagreement and electoral mobilization in intimate networks—a spillover experiment. *American Journal of Political Science*, **61** 289–304.

Haushofer, J. and Shapiro, J. (2016). The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya*. *The Quarterly Journal of Economics*, **131** 1973–2042.

Hidrobo, M., Peterman, A. and Heise, L. (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern ecuador. *American Economic Journal: Applied Economics*, **8** 284–303.

Hirano, K. and Hahn, J. (2010). Design of randomized experiments to measure social interaction effects. *Economics Letters*, **106** 51–53.

References III

Ichino, N. and Schündeln, M. (2012). Deterring or displacing electoral irregularities? spillover effects of observers in a randomized field experiment in ghana. *The Journal of Politics*, **74** 292–307.

Imai, K., Jiang, Z. and Malani, A. (2021). Causal inference with interference and noncompliance in two-stage randomized experiments. *Journal of the American Statistical Association*, **116** 632–644.

Kinnan, C., Malani, A., Voena, A., Conti, G. and Imai, K. (2020). Adverse selection does not explain why utilization rises with premiums: evidence from a health insurance experiment in india. URL `https://extranet.sioe.org/uploads/sioe2020/kinnan_malani_voena_conti_imai.pdf`.

Malani, A., Holtzman, P., Imai, K., Kinnan, C., Miller, M., Swaminathan, S., Voena, A., Woda, B. and Conti, G. (2021). Effect of health insurance in india: A randomized controlled trial. Working Paper 29576, National Bureau of Economic Research. URL `http://www.nber.org/papers/w29576`.

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

References IV

Muralidharan, K. and Sundararaman, V. (2015). The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India *. *The Quarterly Journal of Economics*, **130** 1011–1066.

Rogers, T. and Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour*, **2** 335–342.

Shi, W., Zhao, A. and Liu, H. (2022). Rerandomization and covariate adjustment in split-plot designs. 2209.12385.

Tortarolo, D., Cruces, G. and Vazquez-Bare, G. (2023). Design of partial population experiments with an application to spillovers in tax compliance. IFS Working Papers W23/17, Institute for Fiscal Studies. URL https://ideas.repec.org/p/ifs/ifsewp/23-17.html.

Zhao, A. and Ding, P. (2022). Reconciling design-based and model-based causal inferences for split-plot experiments. *The Annals of Statistics*, **50** 1170 – 1192.

For $(z, h) \in \{(1, \pi_2), (0, \pi_2), (0, 0)\}$,

$$\tilde{Y}_g(z, h) = \frac{N_g}{E[N_g]} \left(\bar{Y}_g(z, h) - \frac{E[\bar{Y}_g(z, h) N_g]}{E[N_g]} \right) ,$$

where

$$\bar{Y}_g(1, \pi_2) = \frac{1}{M_g^1} \sum_{i \in \mathcal{M}_g} Y_{i,g}(1, \pi_2) Z_{i,g}(\pi_2)$$

$$\bar{Y}_g(0, \pi_2) = \frac{1}{M_g^0} \sum_{i \in \mathcal{M}_g} Y_{i,g}(0, \pi_2)(1 - Z_{i,g}(\pi_2))$$

$$\bar{Y}_g(0, 0) = \frac{1}{M_g} \sum_{i \in \mathcal{M}_g} Y_{i,g}(0, 0) \ .$$

Homogeneous partial interference

Potential outcomes satisfy:

$$Y_{i,g}(\mathbf{z}, n) = Y_{i,g}(\mathbf{z}', n) \text{ w.p.1} \quad \text{if } z_{i,g} = z'_{i,g}$$
$$\text{and} \sum_{1 \leq j \leq n} z_{j,g} = \sum_{1 \leq j \leq n} z'_{j,g} \text{ for any } 1 \leq i \leq n, 1 \leq g \leq G \; .$$

where $\mathbf{z}$ and $\mathbf{z}'$ are any realized vectors of assignment, and $z_{i,g}, z'_{i,g}$ are the corresponding individual treatment indicators for i-th unit in g-th cluster.

Assumptions (cont.)

Sampling procedure

The distribution satisfies:

(a) $\{(\mathcal{M}_g, C_g, N_g, X_g, (\bar{Y}_g(z,h), z \in \{0,1\}, h \in \{0, \pi_2\})) : 1 \le g \le G\}$ is an i.i.d. sequence of random variables.

(c) $P\{|\mathcal{M}_g| \ge 2\} = 1$ and $E[N_g^2] < \infty$.

(d) For some constant $C < \infty$,
$P\left\{E[Y_{i,g}^2(z,h) \mid N_g, C_g, X_g] \le C \text{ for all } 1 \le i \le N_g\right\} = 1$ for all $z \in \{0,1\}$ and $h \in \mathcal{H}$ and $1 \le g \le G$.

(e) $\mathcal{M}_g \perp ((Y_{i,g}(z,h) : z \in \{0,1\}, h \in \mathcal{H}) : 1 \le i \le N_g) \mid C_g, N_g, X_g$ for all $1 \le g \le G$.

(f) For all $z \in \{0,1\}, h \in \mathcal{H}$ and $1 \le g \le G$,

$$E\left[\frac{1}{M_g} \sum_{i \in \mathcal{M}_g} Y_{i,g}(z,h) \mid N_g\right] = E\left[\frac{1}{N_g} \sum_{1 \le i \le N_g} Y_{i,g}(z,h) \mid N_g\right] \text{ w.p.1 .}$$

To formalize the second-stage assignment:

$$Z_{i,g} = \sum_{h \in \mathcal{H}} Z_{i,g}(h) I\{H_g = h\} \text{ for } 1 \leq i \leq N_g \ .$$

Second-stage assignment

The treatment assignment mechanism for the second-stage satisfies:

(a) $(((Z_{i,g}(h) : h \in \mathcal{H}) : 1 \leq i \leq N_g) : 1 \leq g \leq G) \perp H^{(G)}$,

(b) $W^{(G)} \perp (((Z_{i,g}(h) : h \in \mathcal{H}) : 1 \leq i \leq N_g) : 1 \leq g \leq G) \mid (B_g : 1 \leq g \leq G)$,

(c) For all $1 \leq g \leq G$,
$E[Z_{i,g}(h) \mid B_g] = \frac{1}{M_g} \sum_{i \in \mathcal{M}_g} Z_{i,g}(h) = h + o_P(1)$.

Lipschitz condition

The distribution of data is such that

(a) $E[\bar{Y}_g^r(z,h)N_g^\ell|S_g = s]$ is Lipschitz in s for $(z,h) \in \{(0,0),(0,\pi_2),(1,\pi_2)\}$ and $r, \ell \in \{0,1,2\}$.

(b) For some $C < \infty$, $P\{E[N_g^2|S_g] \leq C\} = 1$

Variance Estimator

For $z \in \{0,1\}$,

$$\hat{V}(z) = \frac{1}{\pi_1}\hat{\mathbb{V}}^z_{1,n}(\pi_2) + \frac{1}{1-\pi_1}\hat{\mathbb{V}}^z_{1,n}(0) + \hat{\mathbb{V}}^z_{2,n}(\pi_2,\pi_2) + \hat{\mathbb{V}}^z_{2,n}(0,0) - 2\hat{\mathbb{V}}^z_{2,n}(\pi_2,0)$$

with

$$\hat{\mathbb{V}}^z_{1,n}(h) = \hat{\mathbb{E}}\left[\mathrm{Var}\left[\tilde{Y}_g(z,h) \mid S_g\right]\right] := \left(\hat{\sigma}^z_n(h)\right)^2 - \left(\hat{\rho}^z_n(h,h) - (\hat{\Gamma}^z_n(h))^2\right)$$

$$\hat{\mathbb{V}}^z_{2,n}(h,h') = \hat{\mathrm{Cov}}\left[E\left[\tilde{Y}_g(z,h) \mid S_g\right], E\left[\tilde{Y}_g(z,h') \mid S_g\right]\right] := \hat{\rho}^z_n(h,h') - \hat{\Gamma}^z_n(h)\hat{\Gamma}^z_n(h')$$

where

$$\hat{\rho}^z_n(h,h) := \frac{2}{n}\sum_{1\le j\le \lfloor n/2\rfloor}\frac{1}{k^2(h)}\Big(\sum_{i\in\lambda_{2j-1}}\tilde{Y}^z_i I\{H_i = h\}\Big)\Big(\sum_{i\in\lambda_{2j}}\tilde{Y}^z_i I\{H_i = h\}\Big)$$

$$\hat{\rho}^z_n(\pi_2,0) := \frac{1}{n}\sum_{1\le j\le n}\frac{1}{l(k-l)}\Big(\sum_{i\in\lambda_j}\tilde{Y}^z_i I\{H_i = \pi_2\}\Big)\Big(\sum_{i\in\lambda_j}\tilde{Y}^z_i I\{H_i = 0\}\Big)$$

$$\left(\hat{\sigma}^z_n(h)\right)^2 := \frac{1}{nk(h)}\sum_{1\le g\le G}(\tilde{Y}^z_g - \hat{\Gamma}^z_n(h))^2 I\{H_g = h\}\ .$$

Variance Estimator (cont.)

For $z \in \{0, 1\}$,

$$\tilde{Y}_g^z = \frac{N_g}{\frac{1}{G}\sum_{1 \leq g \leq G} N_g} \left(\bar{Y}_g^z - \frac{\frac{1}{G_g}\sum_{1 \leq j \leq G} \bar{Y}_j^z I\{H_g = H_j\} N_j}{\frac{1}{G}\sum_{1 \leq j \leq G} N_j} \right) .$$

where $G_g = \sum_{1 \leq j \leq G} I\{H_g = H_j\}$.

$$\hat{\rho}_n^z(h, h) \xrightarrow{P} E[E[\bar{Y}_g(z, h) \mid S_g]^2]$$

$$\hat{\rho}_n^z(\pi_2, 0) \xrightarrow{P} E[E[\bar{Y}_g(z, \pi_2) \mid S_g] E[\bar{Y}_g(z, 0) \mid S_g]]$$

$$\left(\hat{\sigma}_n^z(h)\right)^2 \xrightarrow{p} \mathrm{Var}[\tilde{Y}_{i,g}(z, h)] .$$

Assumption 3

For $z \in \{0,1\}$, $1 \leq i \neq j \leq N_g$,

$$\begin{aligned}&\operatorname{Cov}\left[Y_{i,g}(z,\pi_2), Y_{j,g}(z,\pi_2) \mid (X_{i,g} : 1 \leq i \leq N_g)\right] \\ &= \operatorname{Cov}\left[Y_{i,g}(z,\pi_2), Y_{j,g}(z,\pi_2)\right] \ .\end{aligned}$$

Linear adjustment for primary effects:

$$\hat{\theta}^{P,adj} = \frac{1}{N_1} \sum_{1 \le g \le G} I\{H_g = \pi_2\}(N_g \bar{Y}_g^1 - (\psi_g - \bar{\psi}_G)' \hat{\beta}^P) - \frac{1}{N_0} \sum_{1 \le g \le G} I\{H_g = 0\}(N_g \bar{Y}_g^1 - (\psi_g - \bar{\psi}_G)' \hat{\beta}^P) \ ,$$

where $\hat{\beta}^P$ is obtained from the linear regression of $\hat{\mu}_{1,j} - \hat{\mu}_{0,j}$ on a constant and $\hat{\psi}_{1,j} - \hat{\psi}_{0,j}$.

$$\hat{\mu}_{1,j} = \frac{1}{k\pi_1} \sum_{g \in \lambda_j} \tilde{Y}_g^1 \bar{N}_g I\{H_g = \pi_2\}$$

$$\hat{\psi}_{1,j} = \frac{1}{k\pi_1} \sum_{g \in \lambda_j} \psi_g I\{H_g = \pi_2\} \ .$$

Under reasonable assumptions, as $G \to \infty$,

$$\sqrt{G}(\hat{\theta}^{P,adj} - \theta^P) \xrightarrow{d} N(0, V^*(1)) \ ,$$

where

$$V^*(1) = V(1) - \frac{1}{\pi_1(1-\pi_1)} \frac{1}{E[N_g]^2} E[\mathrm{Var}[\psi_g' \beta^P \mid S_g]] \ ,$$

with

$$\beta^P = \pi_1(1-\pi_1)(E[\mathrm{Var}[\psi_g \mid S_g]])^{-1} E\left[\mathrm{Cov}\left[\frac{\tilde{Y}_g(1,\pi_2)}{\pi_1} + \frac{\tilde{Y}_g(0,0)}{1-\pi_1}, \psi_g \mid S_g\right]\right] E[N_g] \ .$$