Spillovers

Mauricio Romero

Spillovers

Estimating Treatment Effects with Interference/Spillovers

Issues in Experimental Design for Spillovers

Spillovers

Estimating Treatment Effects with Interference/Spillovers

Issues in Experimental Design for Spillovers

Spillovers

Estimating Treatment Effects with Interference/Spillovers

- How many potential outcomes does each individual have?
- Spillovers within pairs of observations
- Estimating variances in the face of interference
- Partial Population designs
- Randomized Saturation designs
- Problems with standard designs in the face of spillovers
- Weighting in Spillover Experiments

Issues in Experimental Design for Spillovers

- Power
- Gateway to Treatment
- An alternate way of estimating the ToT
- Recovering the Unpolluted Counterfactual in the Absence of a Pure Control

How many potential outcomes does each individual have?

- Standard notation assumes individuals have two potential outcomes
  - Y_{0i} if untreated
  - Y_{1i} if treated
- This assumes away spillover effects
  - General model: as many potential outcomes as there are permutations of treatment among all agents in a network
  - Close analogy with reflection problem (Manski, 1993)

How many potential outcomes does each individual have?

- Individuals who share connections may share similar outcomes
- Impossible using observational data to untangle whether this is because:
  - Share a similar context (contextual effects)
  - Experience similar shocks/covariates (correlated effects)
  - Actual interference between outcomes (endogenous effects)
- Experiments **can** disentangle these effects
- In an experimental framework, spillovers are possible through
  1. Context effects: experiment shifts covariates for the untreated
  2. Direct contagion: Endogenous transmission of outcomes

How many potential outcomes does each individual have?

- Manski (2010) works in a completely general framework that permits arbitrary forms of spillovers between units
- Bridge between observational/IV literature on spillovers and the experimental literature
- As long as spillovers only within reference groups, experiments can identify treatment effects with no further assumptions about the distribution of spillovers
- The **Stable Unit Treatment Value Assumption (SUTVA)** holds across experimental clusters, even if we relax it within clusters

Spillovers

Spillovers within pairs of observations

- Assume treatment may spillover within pairs, but does not spillover across pairs
- This setup induces four potential outcomes:
  - Y_{00} is the pure control outcome
  - Y_{01} is the outcome where only the study unit is treated
  - Y_{10} is the outcome when only the non-study unit is treated
  - Y_{11} is the outcome when both are treated
- In this case we have three causal treatment effects of interest:
  - Average Treatment Effect: $E[Y_{01} - Y_{00}]$
  - Average Spillover on the Non-Treated: $E[Y_{10} - Y_{00}]$
  - Average Spillover on the Treated: $E[Y_{11} - Y_{01}]$

Estimating variances in the face of interference

- Cannot estimate the variance of spillover estimands (and therefore make inferences) without placing additional structure on the problem (Halloran and Struchiner, 1995);(Tchetgen and VanderWeele, 2010)
- Reason: experimental subjects need to serve as potential counterfactuals for each other to estimate sample variance
- Example: Two clusters, 50% of potential agents receive the treatment in each cluster
- Each of these outcomes represents a singleton among all the permutations of potential outcomes that it could have experienced
- Hence, can't infer the variance of treatment/spillover effects by comparing them to each other

Spillovers

Estimating variances in the face of interference

- Assume some variant of the Stratified Interference assumption
- Outcomes respond to the intensity of treatment within a treated cluster, but not to the identity of the individuals who receive treatment
- Consider an individual i located in cluster j
- First, assign a cluster-level saturation of treatment $S_j \in \{0, 1\}$
- Then, assign individual treatment status $T_i \in \{0, 1\}$ conditional on these cluster-level saturations
- Individual compliance with treatment is given by $R_i \in \{0, 1\}$
- X_{ij} is a vector of covariates

Estimating variances in the face of interference

- Assumption: The potential outcome Y_{ij} is a function of T_{ij} and S_j
- Y_{ij} is not affected by the specific identities of the individuals assigned to treatment
- Environment in which we can estimate treatment and spillover effects (by comparing to pure control clusters) and estimate variances (because of Stratified Interference)

Spillovers

Partial Population designs

- A Partial Population design features
  - Some clusters remain entirely untreated
  - Some clusters are partially treated
- Criteria that determines treatment in treated clusters does not need to be randomized
  - As long as it can be established in the control
- The most famous example of a PP experiment is Oportunidades/Progresa from Mexico
  - Eligible and ineligible households within villages that were randomly treated
  - ITT effects by comparing eligible households in treatment versus control villages
  - Spillover effects by comparing ineligible households in treatment versus control villages

Partial Population designs

- Several questions remain unanswered with this design:
  - What would the spillover effects on eligible units be?
  - What are the spillovers that do occur on treated units as a function of the treatment around them?
  - How would the treatment and spillover effects vary if the intensity of treatment was varied at the cluster level?

Spillovers

Randomized Saturation designs

- Two level experiments
  - Randomize the intensity of treatment at cluster level
  - Randomize treatment at the unit level within each cluster

Randomized Saturation designs

- New set of experimental estimands
- **Intention to Treat** (ITT) effect: Difference between individuals offered treatment in a cluster with saturation π and pure control individuals
  - $ITT(\pi) := E(Y_{ic}|T_{ic} = 1; \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0; \pi_c = 0)$
- **Spillover on the Non-Treated** (SNT) effect: Difference between individuals not offered treatment in a cluster with saturation π and pure control individuals
  - $SNT(\pi) := E(Y_{ic}|T_{ic} = 0; \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0; \pi_c = 0)$

- **Total Causal Effect** (TCE): Overall difference between treated and pure control clusters

$$TCE(\pi) \quad = \quad E(Y_{ic}|\pi_c = \pi) - E(Y_{ic}|\pi_c = 0)$$

- **Total Causal Effect** (TCE): Overall difference between treated and pure control clusters

$$\begin{aligned} TCE(\pi) &= E(Y_{ic}|\pi_c = \pi) - E(Y_{ic}|\pi_c = 0) \\ &= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\ &\quad E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \end{aligned}$$

- **Total Causal Effect** (TCE): Overall difference between treated and pure control clusters

$$
\begin{aligned}
TCE(\pi) &= E(Y_{ic}|\pi_c = \pi) - E(Y_{ic}|\pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad \pi E(Y_{ic}|T_{ic} = 0, \pi_c = 0) - (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = 0)
\end{aligned}
$$

- **Total Causal Effect** (TCE): Overall difference between treated and pure control clusters

$$
\begin{aligned}
TCE(\pi) &= E(Y_{ic}|\pi_c = \pi) - E(Y_{ic}|\pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad \pi E(Y_{ic}|T_{ic} = 0, \pi_c = 0) - (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \\
&= \pi \left[E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0, \pi_c = 0)\right] + \\
&\quad (1-\pi) \left[E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0, \pi_c = 0)\right]
\end{aligned}
$$

- **Total Causal Effect** (TCE): Overall difference between treated and pure control clusters

$$
\begin{aligned}
TCE(\pi) &= E(Y_{ic}|\pi_c = \pi) - E(Y_{ic}|\pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \\
&= \pi E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) + (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - \\
&\quad \pi E(Y_{ic}|T_{ic} = 0, \pi_c = 0) - (1-\pi)E(Y_{ic}|T_{ic} = 0, \pi_c = 0) \\
&= \pi \left[E(Y_{ic}|T_{ic} = 1, \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0, \pi_c = 0)\right] + \\
&\quad (1-\pi) \left[E(Y_{ic}|T_{ic} = 0, \pi_c = \pi) - E(Y_{ic}|T_{ic} = 0, \pi_c = 0)\right] \\
&= \pi ITT(\pi) + (1-\pi) SNT(\pi)
\end{aligned}
$$

Randomized Saturation designs

- Individuals treated will experience two treatment effects
  1. A direct treatment effect from the program
  2. Spillover effect that arises from the treatment of other individuals in their cluster
- Decompose the ITT into two components
  - **Treatment on the Uniquely Treated** (TUT): The ITT on a sole individual offered treatment within a cluster (i.e., $ITT(1/n) \approx ITT(0)$)
  - **Spillover on the Treated** (ST) measures the (saturation dependent) spillover effect on individuals treated

$$ST(\pi) := E(Y_{ic}|T_{ic} = 1; \pi_c = \pi) - E(Y_{ic}|T_{ic} = 1; \pi_c = 0)$$

Randomized Saturation designs

$$
\begin{aligned}
ITT(\pi) &= E(Y_{ic}|T_{ic}=1;\pi_c=\pi) - E(Y_{ic}|T_{ic}=0;\pi_c=0) \\
&= E(Y_{ic}|T_{ic}=1;\pi_c=\pi) - E(Y_{ic}|T_{ic}=1;\pi_c=0) + \\
&\quad E(Y_{ic}|T_{ic}=1;\pi_c=0) - E(Y_{ic}|T_{ic}=0;\pi_c=0) \\
&= ST(\pi) + TUT
\end{aligned}
$$

Spillovers

Problems with standard designs in the face of spillovers

- Now consider the effects recovered by standard research designs
- A blocked design features a fixed treatment saturation in each cluster and uses within-cluster controls as counterfactuals
  - In face of SUTVA violations, blocked design is unattractive
  - Control units subject to interference from the treatment
  - Since saturation is fixed (typically at 50%), no way of estimating this interference

Problems with standard designs in the face of spillovers

- The ITT recovered by the blocked design will be:

$$\begin{aligned} ITTB(\overline{\pi}) &= E(Y|T=1;S=\overline{\pi}) - E(Y|T=0;S=\overline{\pi}) \\ &= TUT + ST(\overline{\pi}) - SNT(\overline{\pi}) \end{aligned}$$

- The blocked estimator is biased by the term

$$E(Y|T=0;S=\overline{\pi}) - E(Y|T=0;S=0) = SNT(\overline{\pi})$$

- The blocking leaves no variation in the distribution of π that can be used to investigate this bias term

Problems with standard designs in the face of spillovers

- Clustered design: treatment saturations equal either $\overline{\pi}$ or zero, but no data is gathered on untreated individuals in treatment clusters
- Clustered impacts estimated by comparing treatment and control clusters
  - Recovers the correct ITT as long as SUTVA holds between clusters:

$$\begin{aligned} ITTC(\overline{\pi}) &= E(Y|T=1;S=\overline{\pi}) - E(Y|T=0;S=0) \\ &= r(TUT + ST(\overline{\pi})) + (1-r)SNT(\overline{\pi}) \end{aligned}$$

  - Unable to estimate how the treatment or spillover effects vary with saturations (fixed in a typical clustered design)
  - Cannot differentiate $ST(\pi)$ from $SNT(\pi)$ or $ST(\pi)$ from TUT
  - Instrumenting for uptake with offering imposes SUTVA within clusters as an exclusion restriction
  - Cannot estimate $ToT(\pi)$
  - $ToT(\overline{p}) \neq \frac{ITT(\overline{p})}{r}$ unless $SNT(\overline{p}) = SNC(\overline{p}) = 0$

Problems with standard designs in the face of spillovers

- Blocked designs are at the mercy of spillovers
- Clustered designs provide the correct answer but no way to investigate spillovers
- Partial population designs provide direct evidence of the nature of spillover effects
  - Partial population experiments feature pure control clusters where $S = 0$ as well as partially treated clusters where $S = \pi < 1$
  - In a partial population design data is also gathered on untreated units in treated clusters
  - Such designs provide an experimental estimate of the $SNT(\pi)$ by comparing the within-village controls to the pure controls
  - Cannot shed light on saturation or threshold effects because the treatment saturations in such experiments are either fixed or endogenous

Spillovers

Weighting in Spillover Experiments

- Saturation Weights apply a weight $s_\pi^T = \frac{1}{\pi}$; to treated individuals and a weight $s_\pi^U = \frac{1}{1-\pi}$ to untreated individuals
- For example, a cluster with $\pi = \frac{2}{3}$ has twice as many treated individuals as a cluster with $\pi = \frac{1}{3}$
- Weighting the treated individuals by $s_{2/3}^T = \frac{2}{3}$ and $s_{1/3}^T = \frac{1}{3}$ allows one to calculate a pooled estimate that places equal weight on both clusters, rather than twice as much weight the $\pi = \frac{2}{3}$ clusters

Weighting in Spillover Experiments

- Inverse Propensity Weights are a useful tool for recovering unbiased treatment effect estimates when individual have unequal probabilities of being in specific experimental states
- General rule: experimental observations should be weighted by the inverse of the probability that they had their observed treatment status
  - In a simple experimental design all units have the same probability of treatment. IPWs irrelevant because they are the same for each unit
  - In blocked designs it may be possible that treatment probabilities are not the same within blocks (e.g., a block with three observations and two treatments). IPWs may be necessary to recover unbiased treatment effect estimates
  - In more complex designs, and particularly when analyzing spillovers, IPWs are a transparent way to deal with different probabilities of ending up with certain treatment status

Weighting in Spillover Experiments

- Example: Social network analysis in a population
  - Treat 1 out of every 5 individuals
  - Goal: Understand how spillover effects spread through social networks, but individuals have different numbers of friends
  - Every individual has the same chance of being in the treatment
  - But chance of being in the spillover group (having a treated friend while remaining untreated) depends on the number of friends
  - Individuals with no friends: undefined IPW and no potential outcome for the spillover
  - Parametric way: regression controlling for the number of friends that an individual has. Conditional on this the number of treated friends should be random
  - Economics papers from the beginning of the spillovers literature (e.g., 'Worms' paper) have mostly used this approach
  - A more elegant solution is given in Aranow and Samii (2012), who provide a Randomization Inference-based way of combining IPW weighting with RI to calculate the average treatment (or spillover) effect plus sharp p-values

Spillovers

Estimating Treatment Effects with Interference/Spillovers

Issues in Experimental Design for Spillovers

Spillovers

Estimating Treatment Effects with Interference/Spillovers

Issues in Experimental Design for Spillovers

Spillovers

Power

- Begin by assuming that the data generating process has a random effects error structure
- $\varepsilon_{ic} = v_c + w_{ic}$; with common cluster component $v_c \sim (0, \tau^2)$ and individual component $w_i c \sim (0, \sigma^2)$
- μ: fraction of the sample assigned to the treatment
- ψ: fraction of the clusters assigned to pure control
- η_T^2: variance in treatment saturation within treatment clusters
- C: number of clusters
- n: number of observations per cluster
- A random effects framework combined with randomized saturation decomposes the clustering of outcomes into two components:
  1. Outcomes endogenously driven by treatment of others in the same cluster
  2. The statistical random effect in outcomes, which reduces the power of the clustered estimates but does not imply interference between units

Power

- A simple regression-based estimator of the pooled effects is:

$$Y_{ic} = \beta_0 + \beta_1 T_{ic} + \beta_2 S_{ic} + \phi X_{ic} + \varepsilon_{ic}$$

- Then, given statistical significance level α and power γ, the MDE of ITT_w is:

$$MDE_w^T = (T_{1-\gamma} + t_\alpha)\sqrt{\frac{(n-1)\tau^2\left(\frac{1}{(1-\phi)\phi} + \left(\frac{1-\phi}{\mu^2}\right)\eta_T^2\right) + (\tau^2+\sigma^2)\left(\frac{\phi+\mu}{\mu\phi}\right)}{nC}}$$

- The MDE of $\overline{SNT}_w$ $\left(MDE_w^S\right)$ is similar, substituting μ_s for μ

Power

- There are several distinct sources of power loss from using a RSD:
  1. Need to create within-cluster controls means $\rightarrow$ the total sample size of treated units is smaller than if you had treated 100% of units in treatment clusters
  2. In the presence of non-zero intra-cluster correlation, highest-power design will be a blocked design because all identification is within-cluster
  3. The RSD links the correlation in treatment status with the correlation in outcomes, and hence results in a decrease in power that comes from the product of the ICC and the variation in treatment status across clusters
  4. i.e., you are more likely to have imbalance in the overall experiment when the fraction of treated units varies at the cluster level, if cluster level means are not equal
- **The cost of estimating up these effects is power!**

Power

- If you are trying to maximize the sum of the power of the treatment and spillover effects, then the size of the pure control should be greater than one third (because it is the counterfactual for both groups;).
- In fact greater than .41, but less than half, and the size of the optimal pure control is increasing in the ICC
- The power of the simple impact estimator decreases linearly with the variance in the treatment saturations
- Power to detect a slope term on the saturations increases less than linearly with the variance in the treatment saturations (as the saturations become too spread you have too few of the observations of one treatment status)
- The optimal saturations are symmetric about .5, and are increasingly spread out as the sample size increases

Spillovers

Are you randomizing the saturation of treatment within your study sample, or within the population

- If you are randomizing with a study sample, then the true saturation rate in the population will be the product of the sampling rate and the treatment rate. Is this number large enough to produce credible spillover effects in the population?
- Spillovers will be restricted to occur within your study sample (teenage girls)? In this case the 'true' saturation is defined relative to the target population not the overall population
- If the sampling for eligibility into the study is not random but is based on some (potentially not perfectly) observed criteria, then you can't use the true saturation for the RS analysis but must instead instrument for the (endogenous) true with the (randomized) assigned saturation.

Spillovers

An alternate way of estimating the ToT

- In the face of interference you cannot use the normal strategy of instrumenting for the ToT because the spillovers violate the Exclusion Restriction
- As a possible way around this, assume the average spillover on non-compliers is the same as the average spillover on non-treated
- Then

$$TOC(\pi) = \frac{ITT(\pi) - (1 - r(\pi))(SNC(\pi))}{r(\pi)}$$

An alternate way of estimating the ToT

- To provide estimates of these quantities, we can linearize the system and estimate equation

$$Y_{ic} - \delta_0 + \delta_1 T_{ic} + \delta_2 S_{ic} + \delta_3 (T_{ic} \times \pi_c) + \phi X_{ic} + \varepsilon_{ic}$$

- $\widehat{TOC(\pi)} = \frac{\widehat{\beta_{1\pi}} - (1 - \widehat{r(\pi)})\widehat{\beta_{2\pi}}}{\widehat{r(\pi)}}$
- If we assume the compliance rate is constant with respect to π
- $\widehat{TOC(\pi)} = \frac{\widehat{\beta_1} - (1 - \widehat{r})\widehat{\beta_2}}{\widehat{r}}$
- $\frac{d\widehat{SC(\pi)}}{d\pi} = \frac{\widehat{\delta_3} - (1 - \widehat{r})\widehat{\delta_4}}{\widehat{r}}$
- $\widehat{TUC} = \frac{\widehat{\delta_1}}{\widehat{r}}$
- Cross-equation hypothesis testing can be performed using either Seemingly Unrelated Regression or GMM.

Spillovers

An alternate way of estimating the ToT

- Without a pure control group, we compare treated units in treated clusters to untreated units in treated clusters:

$$Y_{ic} = \beta_0 + \beta_1 T_{ic} + \phi X_{ic} + \varepsilon_{ic}$$

- Without a pure control, counterfactual is at the mercy of within-cluster spillovers
- In this context, the RSD has a different advantage: providing an estimate of the unperturbed counterfactual
- Assuming the response to saturations is continuous at zero, the randomization across the saturations enables a researcher to project what would happen to the counterfactual as the treatment saturation approaches zero
- The following assumption is required: $E(Y|T, \pi)$ is an affine (linear) function of π

An alternate way of estimating the ToT

- Natural to estimate the following saturation regression:

$$Y_{ic} = \delta_0 + \delta_1 T_{ic} + \delta_2 \pi_c + \delta_3 T_{ic} \pi_c + \phi X_{ic} + \varepsilon_{ic}$$

- δ_2: tests whether there is variation in the control outcome across saturations (i.e. spillovers in the control)
- δ_0 is an estimate of $E(Y_{ic} | T_{ic} = 0; \pi_c = 0)$
- β_0 is an estimate of $E(Y_{ic} | T_{ic} = 0; \pi_c > 0)$
- Difference between $\widehat{\beta_0}$ and $\widehat{\delta_0}$ is the $\widehat{SNT_w}$ (the average endogenous effect on controls), which can be used to derive an unbiased estimate of the $\widehat{ITT_w}$

An alternate way of estimating the ToT

- Let ω be a randomized saturation design with no pure control and $\kappa > 2$ interior saturations
- Consistent estimators of $\widehat{ITT_w} = \widehat{\beta}_1 + (\widehat{\beta}_0 - \widehat{\delta}_0)$
- Consistent estimators of $\widehat{SNT_w} = \widehat{\beta}_0 - \widehat{\delta}_0$
- Similar estimates for the ITT and SNT at a specific saturation are generated by estimating equation the model on a single saturation
- This is particularly important for settings in which a pure control is not feasible due to regulatory requirements or other exogenous restrictions
- In McIntosh et al (2014), a Mexican government rule required each participating cluster (municipality) be guaranteed at least one treated sub-unit (neighborhood).