Which of these are causal statements?

1. The girls who went to private school have higher test scores than the girls who went to government school.
2. Inflation went up after the election.
3. Because the weather forecast said it would rain, I carried an umbrella.
4. I carried an umbrella and it rained.
5. Places with higher ethnic diversity have lower public goods provision.

Is the statement true?

1. With random assignment, if the difference in outcomes between Treatment and Control units before the intervention is 0, then the treatment effect is 0.
2. With random assignment, one can obtain an unbiased estimate of the average treatment effect by simple comparison of averages.
3. With random assignment, we can learn individual-level treatment effects.
4. If the treatment effect is 0, the randomization did not work properly.

5. Why is random assignment important?

LECTURE 3: RANDOMIZATION

Nahomi Ichino

EGAP Learning Days – Malawi

14 February 2017

This session

- Review: Random Sampling vs. Random Assignment
- Different designs
  - Access
  - Factorial
  - Timing (aka stepped-wedge)
  - Encouragement
  - Spillover
- Strategies of Randomization
  - Simple
  - Complete
  - Blocked
  - Clustered
  - Factorial
- Essential Good Practices

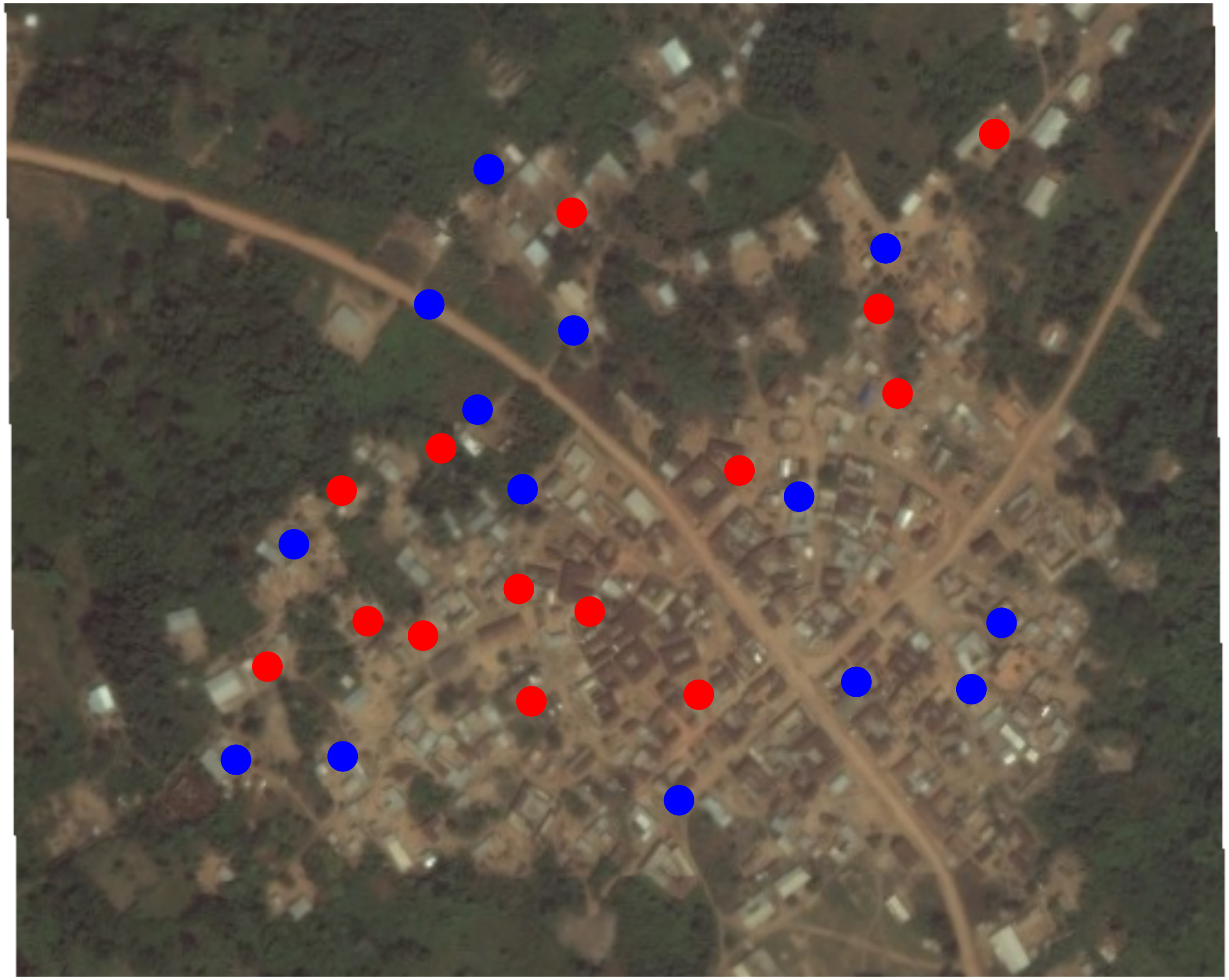
Random Sampling vs. Random Assignment

- Random sampling (*from* population): selecting subjects from a population with known probability

- Random assignment (*to* treatment conditions): assigning subjects with known probability to experimental conditions

Random Sampling

Random Assignment



Random assignment as random sampling from potential outcomes

Random assignment as random sampling from potential outcomes

Strict Definition of Random Assignment

- Every observation must have the same known probability between 0 and 1 of being placed in the treatment group.

Designs

1. Access
2. Factorial
3. Waitlist (aka stepped-wedge)
4. Encouragement
5. Two-Level

Design I - Access

- Lottery: When you don't have enough resources to treat everyone, randomly select a treatment group
- This randomizes **access** to the program
- Example: Merit scholarship program in Kenyan primary schools
  - Positive impact on academic performance of girls eligible for the program in the treatment schools
  - In one district, positive impact on academic performance of boys (ineligible) in the treatment schools
  - Improve teacher attendance in treatment schools
  - Michael Kremer, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn," *Review of Economics and Statistics* 91(3): 437-456.

Design I - Access

- Sometimes, some units (peoples, communities) must have access to a program.
  - EXAMPLE: a partner organization doesn't want to risk a vulnerable community NOT getting a program (want a guarantee that they will be always be treated).

- You can exclude those units, and do random assignment among the remaining units that have a probability of assignment strictly between (and not including) 0 and 1.

Design II - Factorial

- Factorial design enables testing of more than one treatment
- You can analyze one treatment at a time
- Low risk

	T2=0	T2=1
T1=0	25%	25%
T1=1	25%	25%

Design II - Factorial

MONITORING CORRUPTION 207

TABLE 1
NUMBER OF VILLAGES IN EACH TREATMENT CATEGORY

	Control	Invitations	Invitations Plus Comment Forms	Total
Control	114	105	106	325
Audit	93	94	96	283
Total	207	199	202	608

NOTE.—Tabulations are taken from results of the randomization. Each subdistrict faced a 48 percent chance of being randomized into the audit treatment. Each village faced a 33 percent chance of being randomized into the invitations treatment and a 33 percent chance of being randomized into the invitations plus comment forms treatment. The randomization into audits was independent of the randomization into invitations or invitations plus comment forms.

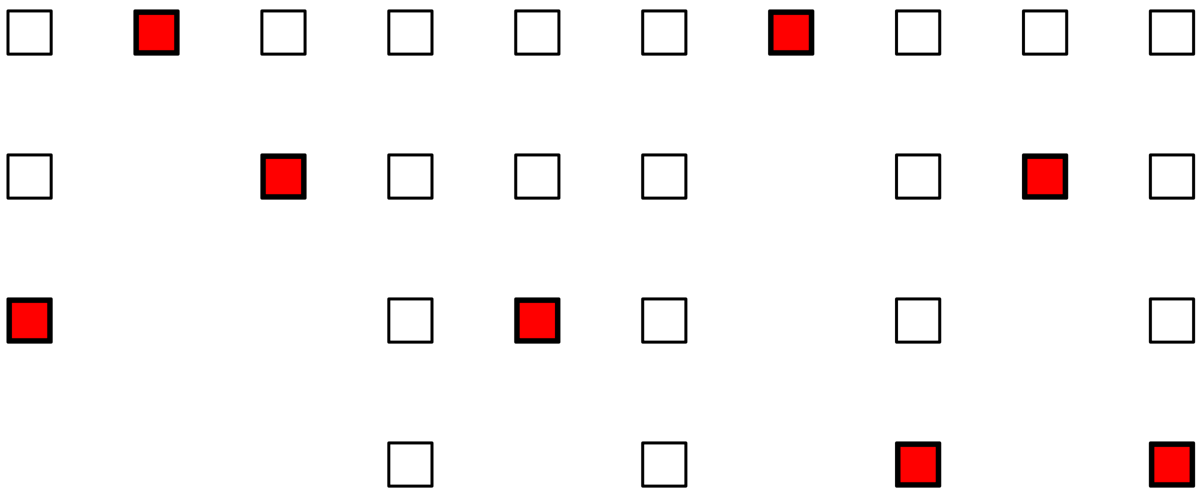
- Olken (2007) Monitoring Corruption: Evidence from a Field Experiment in Indonesia, *Journal of Political Economy*, volume 115: 200 - 249

Design III –Waitlist (aka stepped wedge)

- Randomize **timing of access** to the program
- When an intervention can be or must be rolled out in stages, you can randomize the order in which units are treated
- Can be useful if you don't have the capacity to implement the treatment in a lot of places at once.
- EXAMPLE: a deworming program that is implemented in different schools in stages

Design III –Waitlist (aka stepped wedge)

- Your control group are the as-yet untreated units
  - Be careful: the probability of assignment to treatment will vary over time



Design IV – Encouragement

- Randomizes invitations to subjects to participate in a program.

- Useful when you cannot force a subject to participate and a program is ONLY available through the invitation.
  - Instrumental variables, exclusion restriction
  - Vouchers for private school, attending private school, academic performance

- We can learn the complier average causal effect:

  the causal effect of the participation (not the invitation!) for the units that participate when invited and don't participate when not invited.

Spillovers

- When one subject responds to another subject's treatment status, the non-interference assumption is violated
  - EXAMPLE: vaccinations and herd immunity
  - EXAMPLE: people discuss what they have learned in an education program with their neighbors who did not go to the program

- What is the problem if we randomize treatment assignment to units, and we estimate the Average Treatment Effect as:
  - Mean(Ys for the treated units) – Mean (Ys for the control units)?

Spillovers

- One problem is that $Y_i(1)$ and $Y_i(0)$ can be different (not stable) depending on which other units are treated, so $Y_i(1)$-$Y_i(0)$ is not well-defined.

- One example: Y = days I am sick this year (my vaccination status, my roommate's vaccination status)

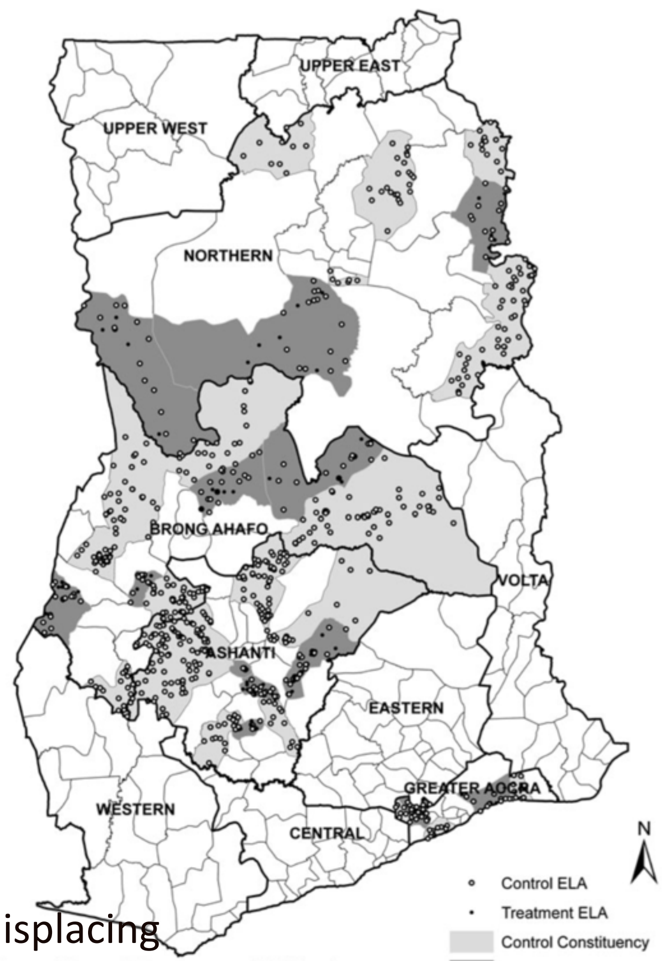
Y(0,0)	30
Y(0,1)	20
Y(1,0)	10
Y(1,1)	5

Spillovers

- You might be able to mitigate the problem:
- If you're worried about spillovers within a community because people talk to each other within a community, randomize at the community level (cluster randomize).
- If you're worried about spillovers across communities because people visit each other across communities, sample your communities (before treatment assignment) in a way that guarantees that the communities are far apart.
- But sometimes, you want to study the spillovers.

Design V – Two-Level

- CODEO's observers for voter registration in 2008
- Two-level design
  - Constituency
  - Registration Centers
- Y(0,0), Y(1,0), Y(1,1)



Ichino, Nahomi, and Matthias Schündeln. 2012. Deterring or displacing electoral irregularities? Spillover effects of observers in a randomized field experiment in Ghana. Journal of Politics 74(1): 292-307.

Control group

- What type of control group is needed?
  - No intervention?
  - Placebo intervention?

- EXAMPLE: Community meetings to listen to drama with different messages
  - Do you want to learn the effect of community meetings + content of drama?
  - Do you want to learn the effect of the content of the drama, given that there are community meetings?
  - E.L.Paluck and D.P.Green, 2009, "Deference, Dissent, and Dispute Resolution: An Experimental Investigation using Mass Media to Change Norms and Behavior in Rwanda." *American Political Science Review* 103: 622-44.

Random Assignment to Relevant Units

- Treatment can be assigned at many different levels: individuals, groups, institutions, communities, time periods, or many different levels.

- You may be constrained in what level you can assign treatment and measure outcomes.

- Your choice of analytic level affects what your study can demonstrate.

Randomization strategies

- Simple
- Complete
- Block
- Cluster
- Factorial

1- Simple

- For each unit, flip a coin to see if it will be treated. Then you measure outcomes at the coin-level.
- The coins don't have to be fair (50-50), but you have to know the probability of treatment assignment.
- You can't guarantee a specific number of treated units and control units.
  - EXAMPLE: If you have 6 units and you flip a fair coin for each, you have about a 3% chance of getting all units assigned to treatment or all units assigned to control.
  - $(1/2)^6 + (1/2)^6$

2- Complete

- A fixed number *m* out of *N* units are assigned to treatment.
- The probability a unit is assigned to treatment is *m/N*. The number of ways treatment can be assigned is N!/(m!*(N-m)!).
- Sometimes we want to make sure that some units of a certain type get treated. This is not guaranteed by complete randomization.
  - EXAMPLE: say I want to assign 4 out of 8 communities to treatment, and I have 4 communities each from 2 different districts
  - The probability that treatment will be assigned to communities in just one district is 1/35.

2- Complete

Gender	Random n	Rank	Select?
F	0.1011	7	1
F	0.3943	5	1
F	0.6757	3	0
F	0.0184	8	1
M	0.2660	6	1
M	0.9889	1	0
M	0.7971	2	0
M	0.5499	4	0
Average			0.5

Done by computer
Simply give a random number to each of *N* units
Then select the *m* units with the highest random number

3- Block

- We can create blocks of that category and randomize separately within each block. You are doing mini-experiments in each block.
  - EXAMPLE: block= district, units= communities

- Probability of treatment assignment can be different in each block, but you have to be very careful about this.
  - EXAMPLE: in the first district, assign 2 out of 6 communities to treatment, and in the second district, assign 4 out of 13 communities to treatment

3- Block

Gender	Block	Random number	Rank	Select?
F	1	0.1378	4	1
F	1	0.4557	3	1
F	1	0.4660	2	0
F	1	0.7909	1	0
M	2	0.9317	1	0
M	2	0.2312	4	1
M	2	0.3993	3	1
M	2	0.9291	2	0
Average				0.5

3- Block

- Advantages to blocking on features that predict the outcome:
  - Guarantee that units of every "type" get some treatment, so treatment and control groups are more similar distributions of these types than without blocking
  - If the blocks are large enough: you can estimate treatment effects for those subgroups
  - Usually improves *power* – your probability of detecting a treatment effect if there is one
- Generally, block if you can.

4- Cluster

- A cluster is a group of units, and all units in the cluster get the same treatment status. This is assigning treatment at the cluster-level.

- Use if the intervention has to work at the cluster level.
  - EXAMPLE: teacher incentives program on student achievement randomized at the school level.

- Having fewer clusters hurts your power. How much depends on the intra-cluster correlation (rho).
  - Higher is worse.

Low intra-cluster correlation (rho)

High intra-cluster correlation (rho)

4- Cluster

City	Cluster	Random n	Rank	Select?
A	1	0.1993	3	1
A	1			1
B	2	0.3836	2	0
B	2			0
C	3	0.1247	4	1
C	3			1
D	4	0.4267	1	0
D	4			0
Average				0.5

Done by computer
Simply give a random number to each of *N* CLUSTERS
Then select the *T* CLUSTERS with the highest random number

4- Cluster

- For the same number of units, having more clusters and smaller clusters can help. But don't go so small that spillovers become a concern.
  - EXAMPLE: teacher incentives program on student achievement randomized at the teacher level.

- Be careful if your clusters have different number of units.
  - EXAMPLE: cluster = set of younger siblings of an adolescent; T = adolescent gets training

You can combine blocking and clustering

	Cluster 1	Cluster 2
Block 1	T	C
Block 2	T	C
Block 3	C	T
Block 4	T	C

- Treatment =durbar
- Unit = individuals
- Cluster = community
- Blocks = communities with similar characteristics

5- Factorial (with blocks)

Gender	Block	Random n	Rank	T1	T2
F	1	0.0444	4	1	1
F	1	0.8061	2	0	1
F	1	0.0660	3	1	0
F	1	0.9680	1	0	0
M	2	0.5482	2	0	1
M	2	0.9003	1	0	0
M	2	0.0784	4	1	1
M	2	0.2565	3	1	0
Average				0.5	0.5

Essential good practices 1

- Verify that the random assignment was implemented properly

- Make sure your random assignment is replicable
  - Set a seed – don't use Excel
  - Preserve your distribution

- Check overall balance with an F-test (treatment assignment on LHS and covariates on RHS)
  - Random assignment gives us, in expectation, overall balance on the covariates.
  - You will see t-tests of covariates one by one. Just by chance, you might get differences on one variable.

Essential good practices 2

- After random assignment, don’t make the T and C groups different by treating them differently!
  - Don’t take extra measurements of the T group
  - Use the same measurement strategy
  - As much as possible, other people should not know whether a unit is in T or C so they don’t treat them differently
  - Enumerators don’t need to know whether they are surveying a T or C community

This session

- Review: Random Sampling vs. Random Assignment
- Different designs
    - Access
    - Factorial
    - Timing (aka stepped-wedge)
    - Encouragement
    - Two-Level
- Strategies of Randomization
    - Simple
    - Complete
    - Blocked
    - Clustered
    - Factorial
- Essential Good Practices

Credits

- A few slides are from Don Green from his book with Alan Gerber, *Field Experiments*
- Many slides are adapted from previous workshops