Lecture 3: Randomization

Maarten Voors and EGAP Learning Days Instructors

9 April 2019 — Bogotá

Learning Days X

Today

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

Exercise

- See handout
- 15mins work in pairs / groups of three

- Make notes for yourself

- 10min plenary discussion

Recap these key terms (more tomorrow)

- Sampling distributions
- Standard deviation (and variation)
- Standard error
- Confidence interval
- Central limit theorem
- p-value
- T-test

Three core assumptions

1. Random assignment of subjects to treatments
   - implies that receiving the treatment is statistically independent of subjects' potential outcomes

Three core assumptions

2. Non-interference: a subject's potential outcomes reflect only whether they receive the treatment themselves
    - So unaffected by how the treatments happened to be allocated
    - i.e. there are no spillovers
    - or SUTVA holds (stable unit treatment value assumption)

Three core assumptions

3. Excludability: a subject's potential outcomes respond only to the defined treatment, not other extraneous factors that may be correlated with treatment
   - Importance of defining the treatment precisely
   - Maintaining symmetry between treatment and control groups (e.g., through blinding, behavioral measures, etc)
   - No attrition

Absent from the list of core assumptions…

- Random sampling of subjects from a larger population is not a core assumption
    - Though random assignment is like random sampling from two alternative universes

- The issue of external validity is a separate question that relates to the issue of whether the results obtained from a given experiment apply to other subjects, treatments, contexts, and outcomes

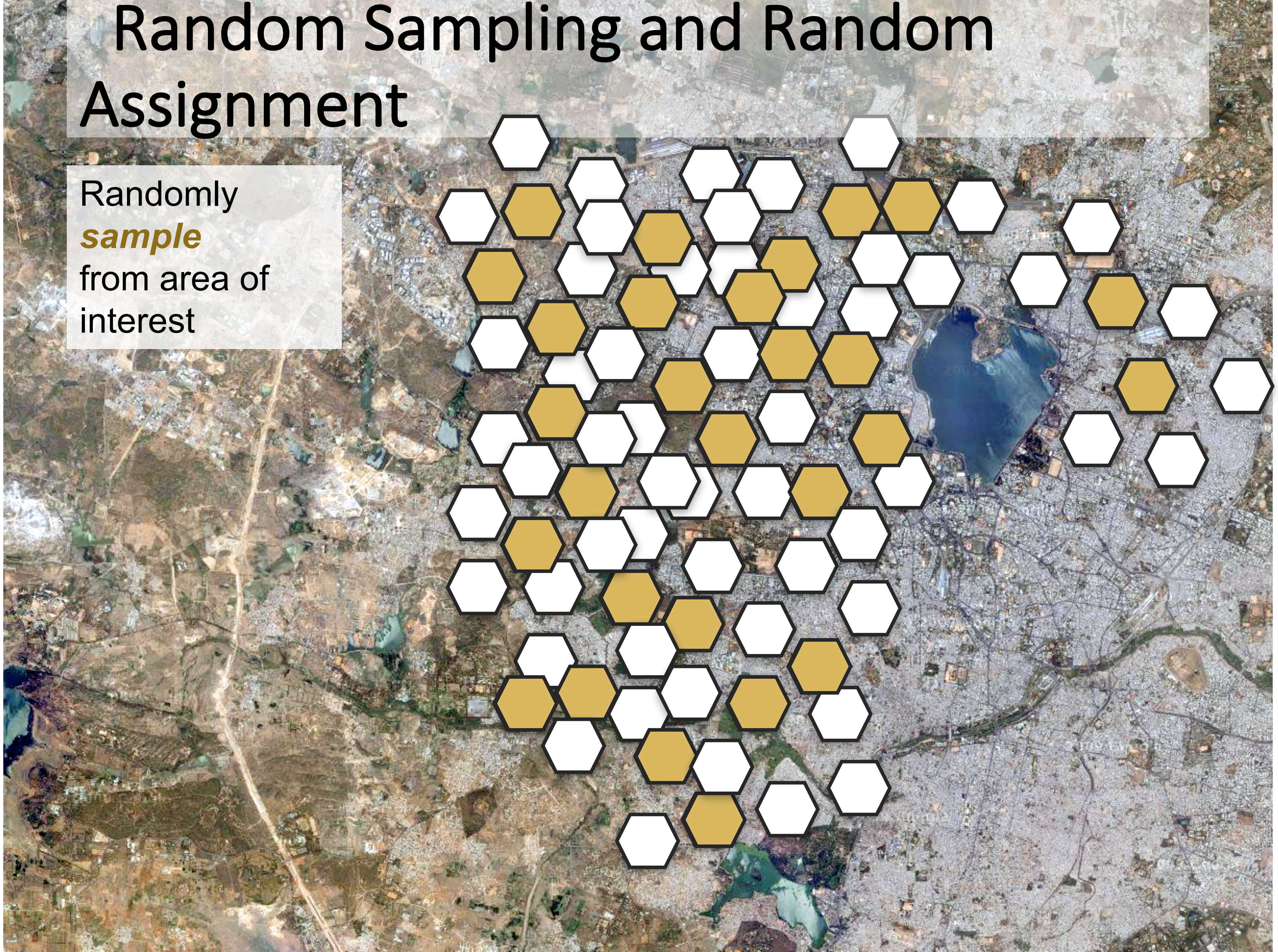
Random Sampling vs. Random Assignment

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

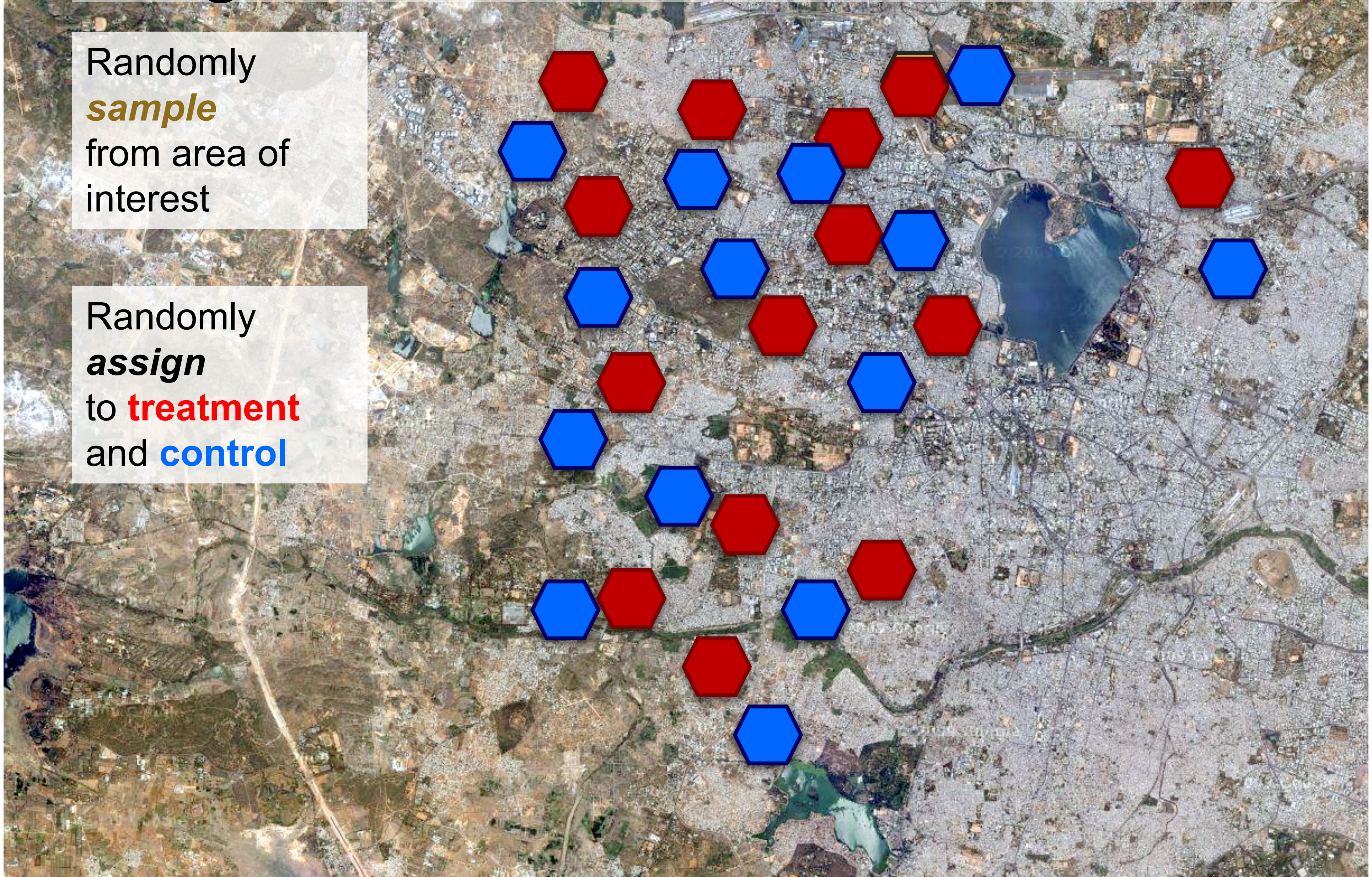
- Random assignment (*to* treatment conditions): assigning subjects with known probability to experimental conditions

Random Sampling and Random Assignment

Randomly ***sample*** from area of interest



Random Sampling and Random Assignment



Randomly ***sample*** from area of interest

Randomly ***assign*** to **treatment** and **control**

Strict Definition of Random Assignment

- Every observation must have the same known probability
- between 0 and 1

Randomization Designs

1. Access

2. Factorial

3. Waitlist (aka stepped-wedge)

4. Encouragement

Randomization Design I - Access

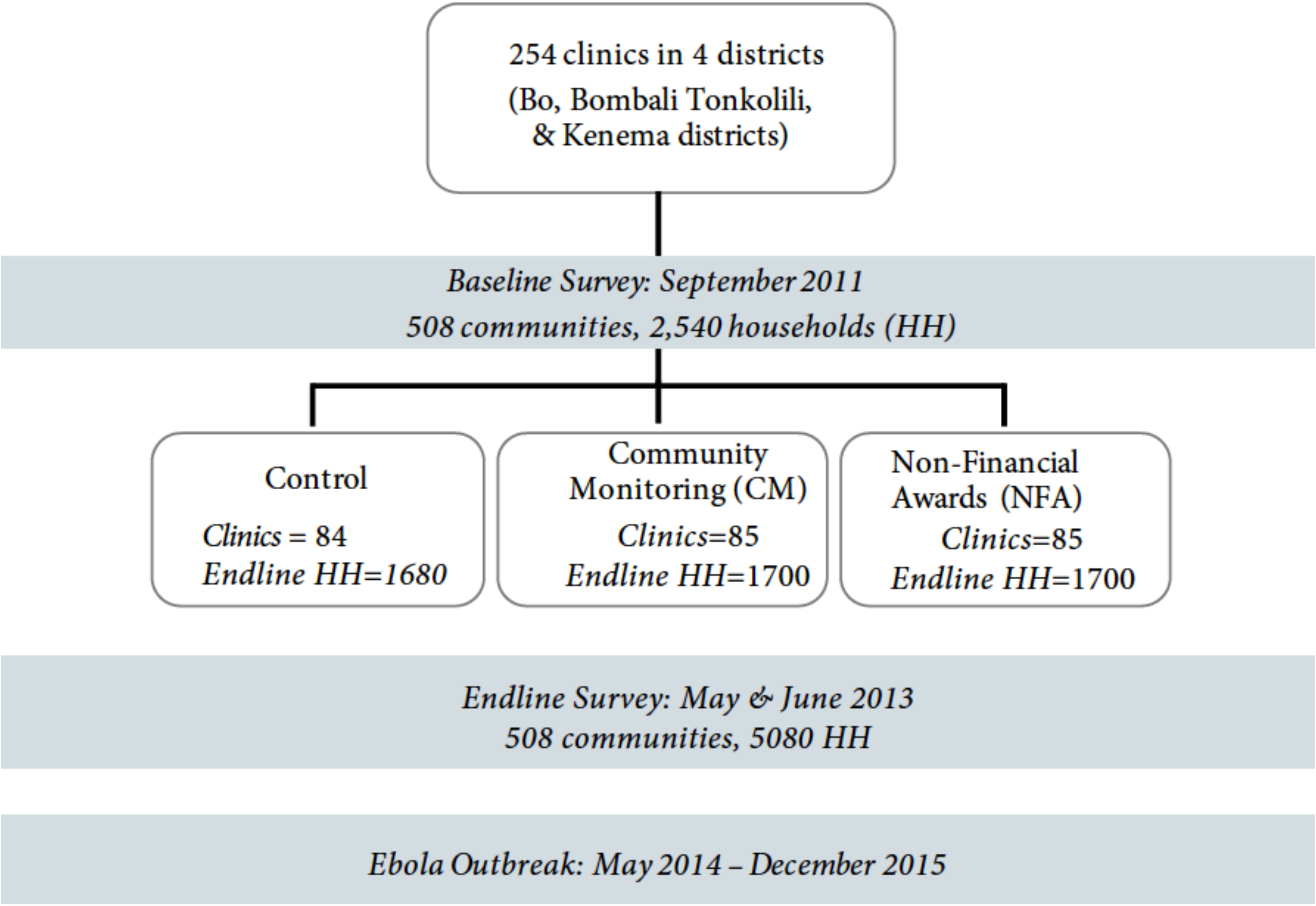
- Through a lottery

- For example ,when you do not have enough resources to treat everyone, randomly select a treatment group

- This randomizes **access** to the program

- Example: Health interventions in Sierra Leone

Consort Diagram



Randomization 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.

Randomization Design II: Factorial Design

- Factorial design enables testing of more than one treatment
- You can analyze one treatment at a time
- Or combinations thereof

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

Example: Colombian Bureaucrats (Tara)

- Phone audit on bureaucrats administering social programs (SISBÉN and Más Familias en Acción) in Colombian alcaldías
- High dimensional, two dimensions were:
  - Social class
  - Regional accent

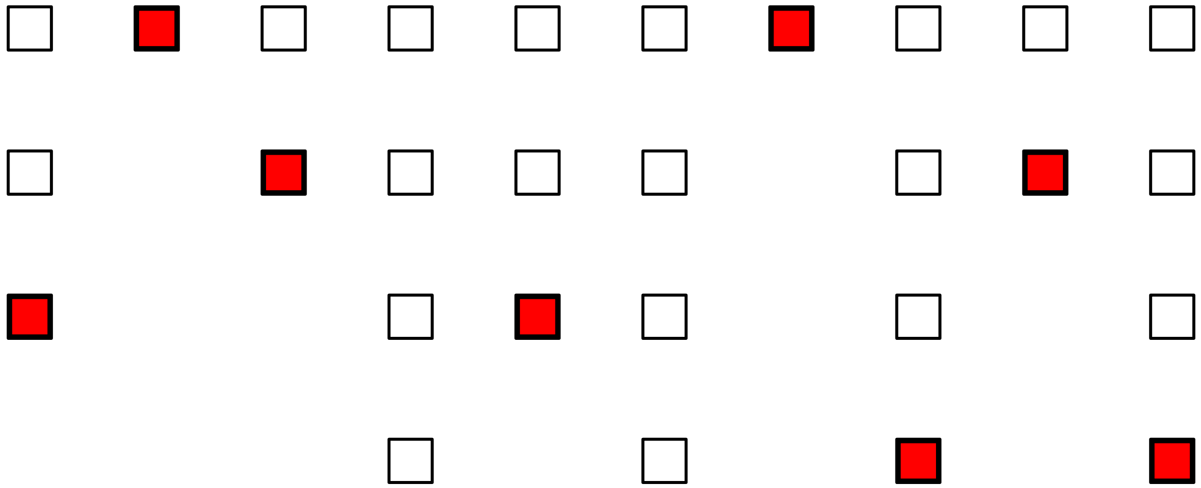
	Bogotá	Costeño	Paisa
Lower Class (~ estratos 1 and 2)	306 calls	306 calls	306 calls
Lower Middle Class (~ estrato 3)	306 calls	306 calls	306 calls

Randomization Design III -Timing of access

- 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
- Often you do not the capacity to implement the treatment in a lot of places at once.

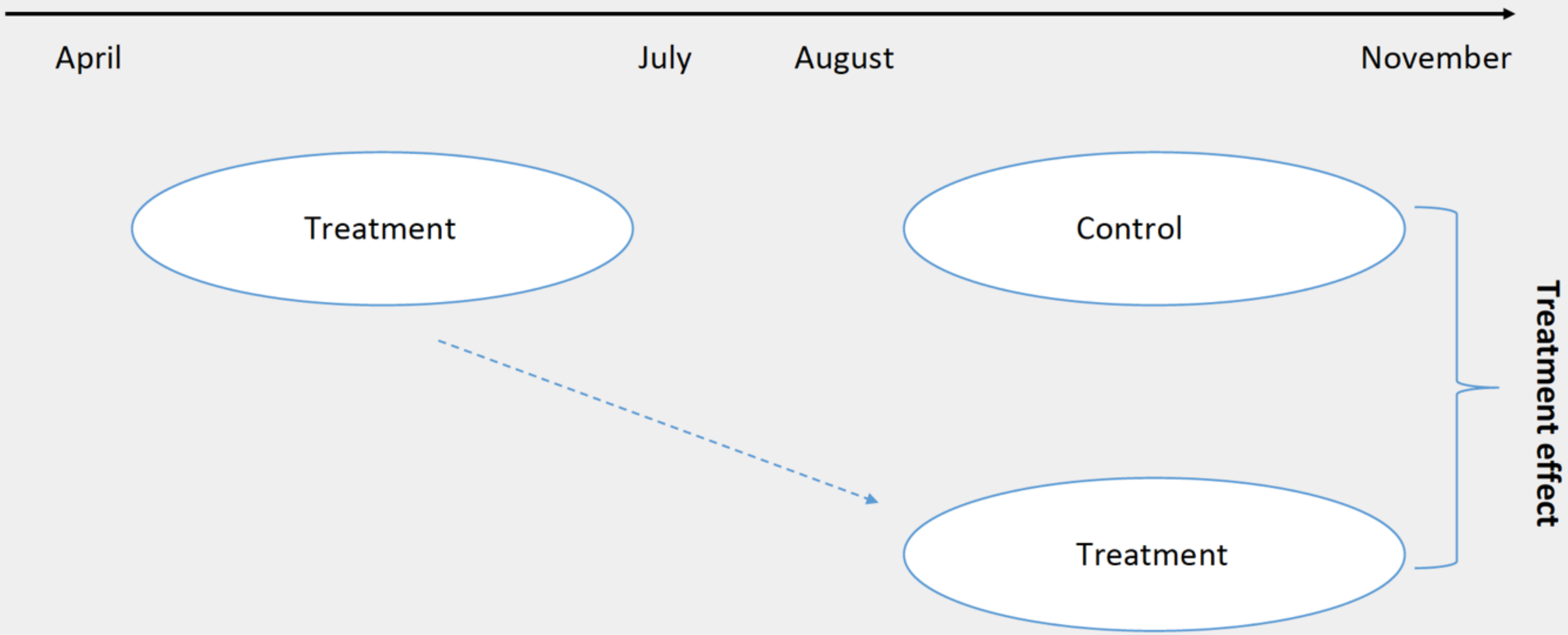
Randomization Design III -Timing of access

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



Research Design

60 eligible municipalities,
All should be treated



Vargas EGAP presentation: The Twists and Turns in the Road to Justice and Peace in Colombia

Randomization Design IV - Encouragement design

- 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 average causal effect for compliers:

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

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.
- Your design?

Control groups

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

- Example:
  - Did a new Hausa television station in northern Nigeria change attitudes about violence, the role of women in society, or the role of youth in society?

- Do you want to learn the effect of watching a film + content of drama?
- Do you want to learn the effect of the content of the drama, given that people are watching a film?
- Or both?

Implementing randomization designs

1. Simple
2. Complete
3. Cluster
4. Block
5. Factorial

- With a computer in advance! (if you can)

Basic Randomization

- Excel
- Stata
- R

Simple Randomization

- 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.
  - $(1/2)^6 + (1/2)^6$

Example

- Excel
- Stata
- R (in a bit)

Complete Randomization

- Most cases

- A fixed number m out of N units are assigned to treatment.

- The probability a unit is assigned to treatment is m/N.

Complete Randomization

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

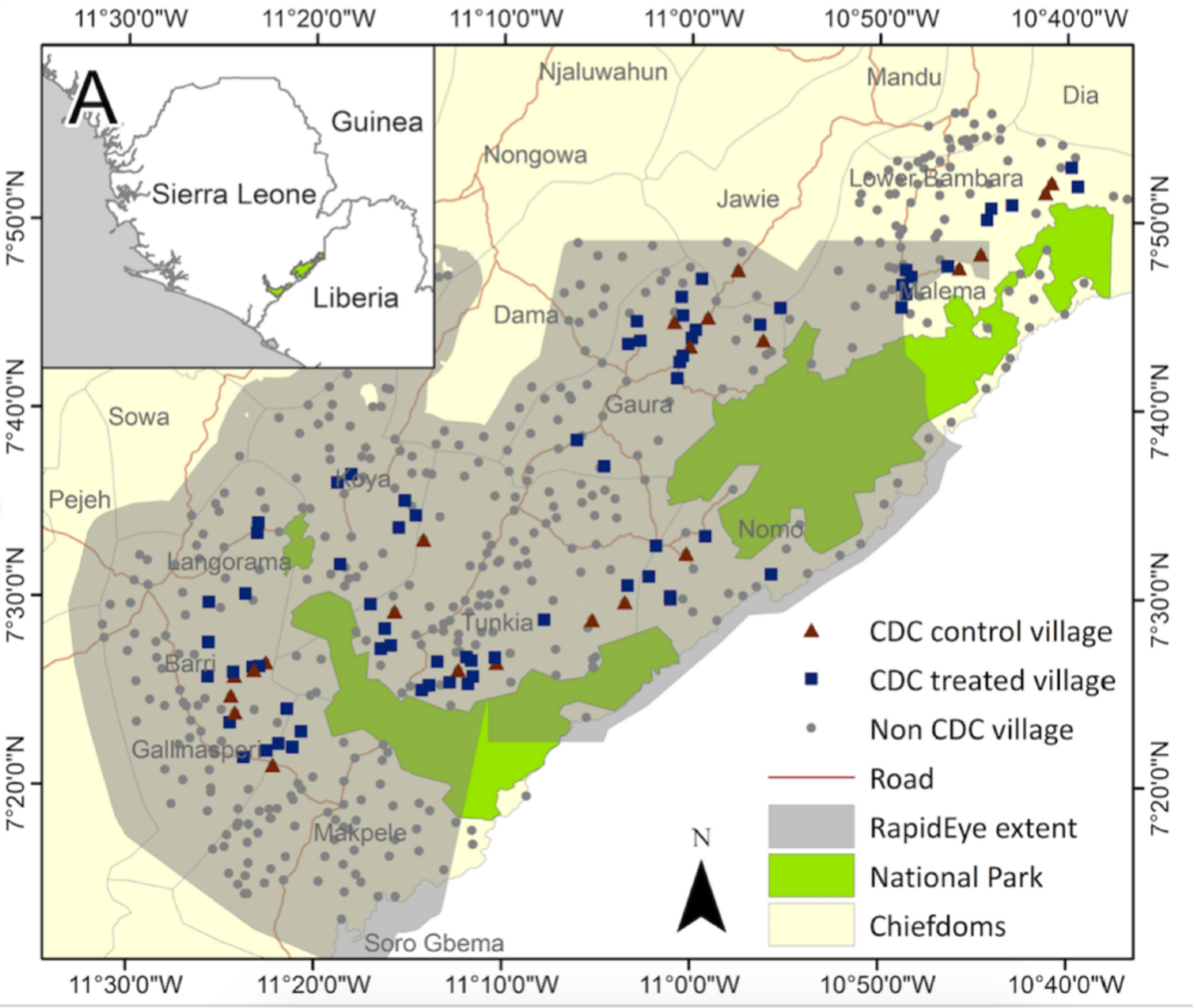
Done by computer
Simply give a random number to each of N units
Then select the T units with the highest random number

Block randomization

- 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
  - Example: Unconditional Transfers and Deforestation
  - Blocks: Chiefdoms, n_j= 6
  - Villages: n = 68

- 68 villages
- 46 aid
- 22 no aid



Block randomization

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

Block randomization

- Advantages to blocking on features that predict the outcome:
  - Guarantee that some units of every “type” get treatment,
  - 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.

Cluster randomization

- 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.

Knowledge

VARIABLE	(1) Knowledge of Law 1448 (No. Correct Answers)	(2) Answered "don't know" in knowledge questions	(3) Do you know about the Victims' Law? (Self-Assessed)	(4) Knows a proper place to make a complaint	(5) Knowledge Standardized index
Treatment	-0.1671 (0.1826)	0.0296 (0.0565)	-0.0496 (0.0530)	0.0483 (0.0499)	-0.0737 (0.1770)
Pair FE	Yes	Yes	Yes	Yes	Yes
Interviewer FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Control Mean	2.679	0.140	0.533	0.692	0.188
Observations	1,054	1,054	3,829	3,829	3,829
R-squared	0.1575	0.1139	0.1245	0.1076	0.0909

				-	N/A
				-	N/A
				-	N/A

Individual victim-level inference

- Simple mean comparison by OLS

$$Y_{impt} = \alpha_p + \rho_t + \beta_1 Treatment_{mp} + \beta_2' X_i + \beta_3' X_{mp} + \varepsilon_{impt}$$

Cluster randomization

- 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: Vargas’ study. Clusters are the towns, units of analysis are people

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

Cluster randomization

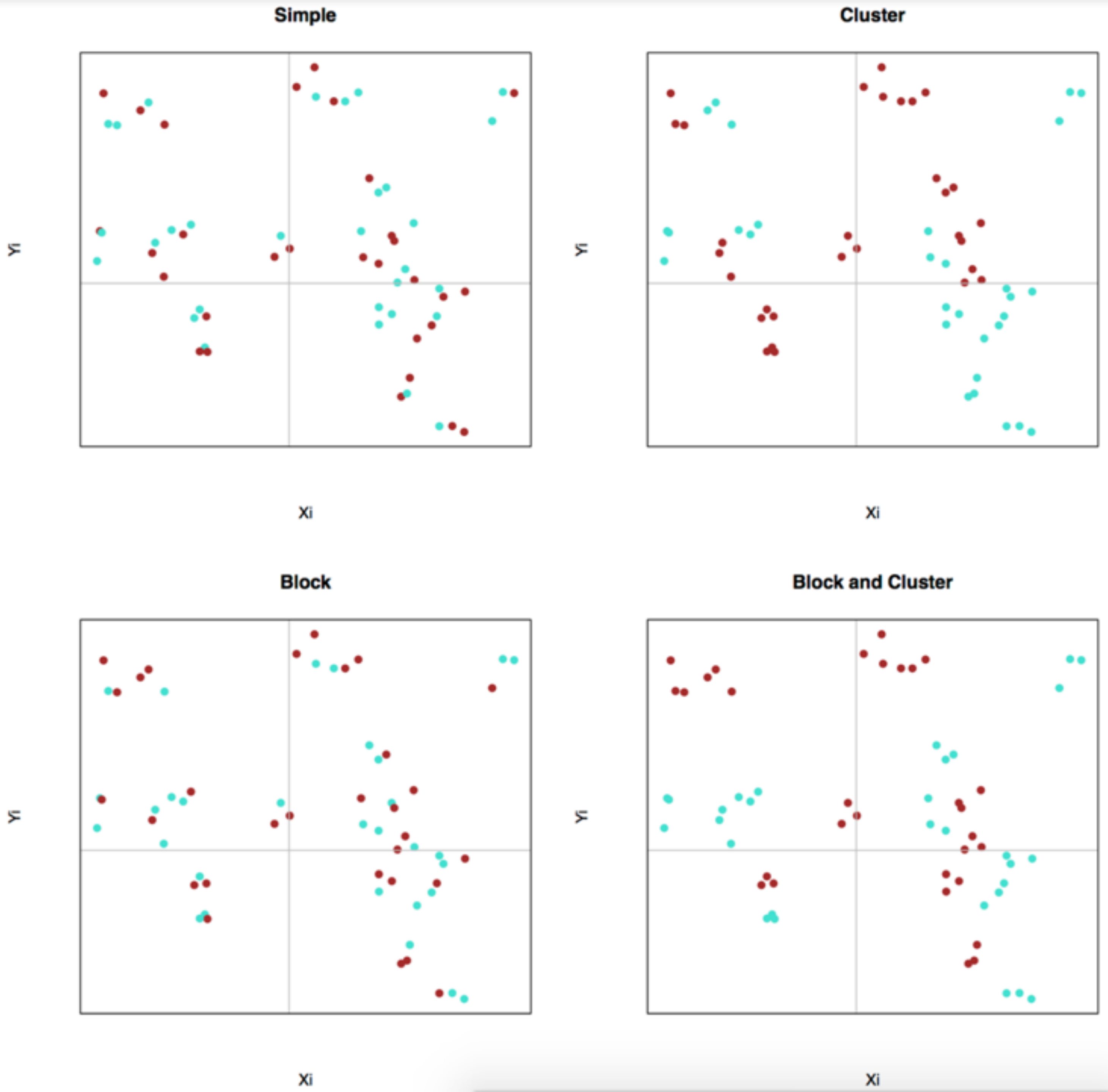
City	Cluster	Random n	Rank	Select?
A	1	0.1993	3	1
A	1			1
B	2	0.3836	2	0
B	2			0
C	3	0.1247	4	1
C	3			1
D	4	0.4267	1	0
D	4			0
Average				0.5

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

Cluster randomization

- For the same number of units, having more clusters and smaller clusters can help.

- Trade off spillover and power



Simple
Cluster
Block
Block and Cluster
Yi
Xi

Did Randomization Work?

- Of course: always

- Make it replicable – Set a seed!
  - Don’t use excel

- Sometimes increased transparency > replicability
- Preserve distributions
- Verify

Good practice

- Check overall balance with an D-square test (Hansen and Bowers 2008)
- Use 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.

Good practice

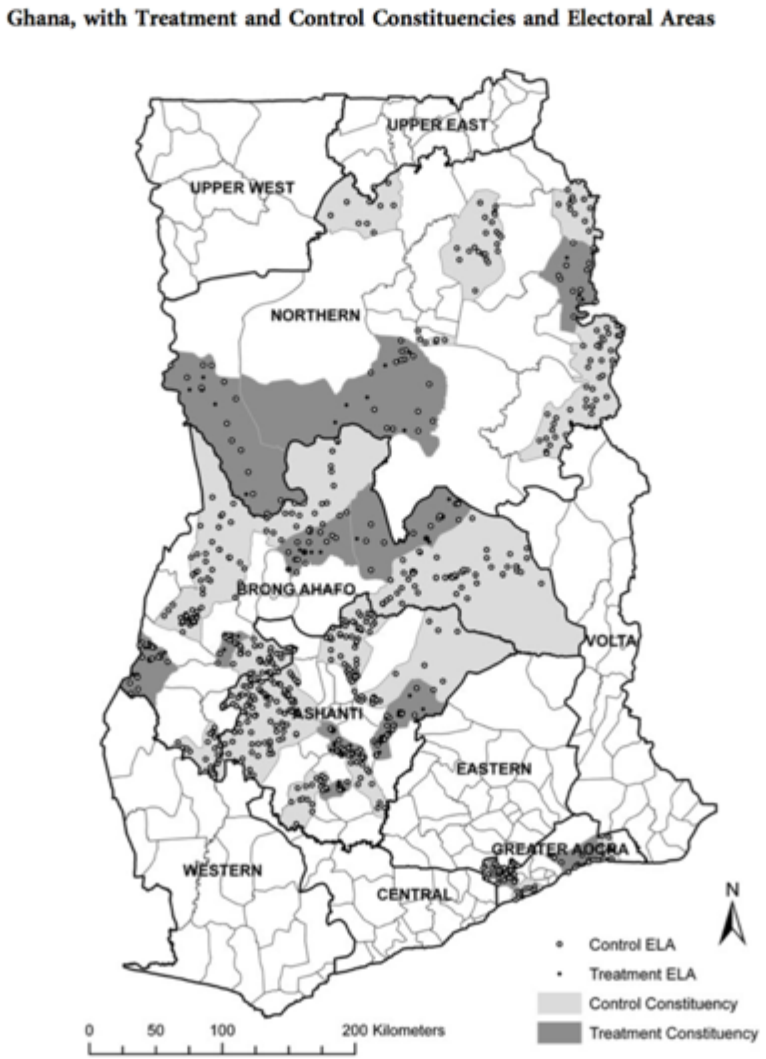
- After random assignment, don't make the T and C groups different by treating them differently! – maintain symmetry
  - Don't take extra measurements of the T group
  - Use the same measurement strategy
  - As much as possible, other people should not know whether a unit is in T or C so they don't treat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

Levels of randomization II

Randomize in such a way as to minimize **spillovers**

- When one subject learns of another subject's treatment status, the non-interference assumption is violated
- The non-interference assumption states that a subject can receive the treatment *only if* he or she is assigned to treatment status
  - e.g. neighbors discuss what they have learned in an education program with their neighbors
- Treating at the community level can sometimes mitigate this

- Randomization for Spillovers
- Two level designs
  - Control
  - Spillover control
  - Treatment



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.

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:

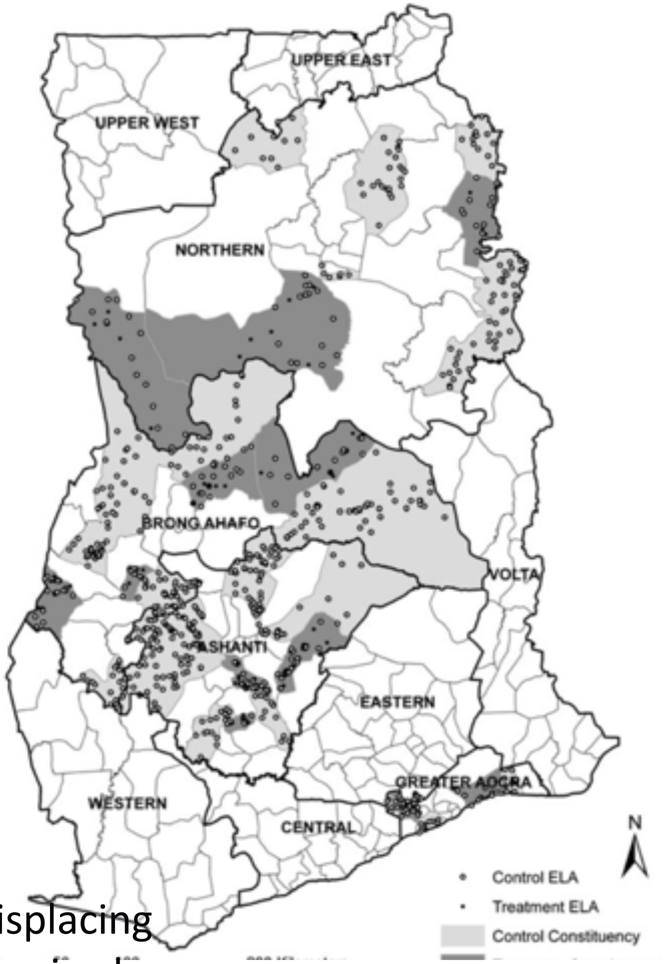
- If you’re worried about spillovers within a community because people talk to each other within a community, randomize at the community level (cluster randomize).

- 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.

Nine Limitations of Randomization

1. Ethics – is this sort of manipulation ethical? Sometimes not.
2. The *real time* constraint. Sometimes to slow. Not much good to help understand history
3. The problem of cost (sometimes; but possible very low)
4. The power constraint. You need a lot of units
   (actually: a problem for any statistical approaches)
5. External validity (problem for any evaluation)
6. The problem of spillovers (problem for any evaluation)
7. The *variables as attributes* constraint (problem for any evaluation)
8. The *assignment to treatment* constraint.
9. Reduced flexibility for organization (problem for any prospective evaluation)

Disclaimer

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

- Many slides are adapted from previous lectures and EGAP workshops