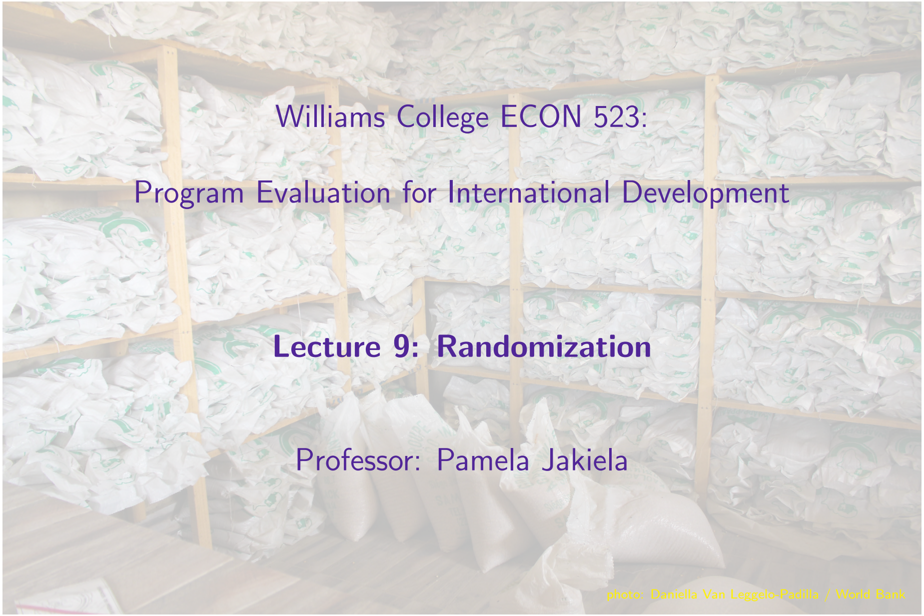


Williams College ECON 523:

Program Evaluation for International Development

Lecture 9: Randomization

Professor: Pamela Jakiela

photo: Daniella Van Leggelo-Padilla / World Bank

The Mechanics of Random Assignment

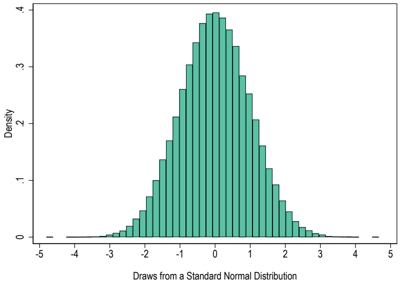
Random Assignment in Three Steps

Randomly assigning treatment using a computer program involves three steps:

1. Assign each eligible unit a (pseudo-)random number
2. Sort the units based on the random numbers
3. Assign treatment based on the random sort order

Step 1: Assign Each Eligible Unit a Random Number

```
clear all
use names.dta
gen rand1 = rnormal()
```



Step 2: Sort Units by Random Number

```
sort rand1
```

Step 3: Assign Treatment Based on Random Sort Order

```
egen treatment = seq(), from(0) to(1)
```

Making Treatment Assignments Reproducible

Start every treatment assignment do file by setting the seed: `set seed 314159`

Randomizing at the Right Level

SUTVA

The **Stable Unit Treatment Value Assumption (SUTVA)**:
Potential outcomes of individual i do not depend on another unit j's treatment assignment

When is SUTVA likely to be violated?

- SUTVA violations matter when we anticipate detectable spillovers unto other eligible units

Conceptually, you should design your RCT so that you are randomizing at a high enough scale to avoid (serious) SUTVA violations (e.g. at the village or school rather than the child level)

Cluster-Randomized Trials

An RCT is **cluster-randomized** if treatment assignment occurs at a higher level than outcomes

- Example: extension agent assigned to village, but farmer-level data on inputs and crops

When to cluster, and how to choose the level at which to assign treatments:

- You cannot assign treatments at a level below the level of data collection
- You should not assign treatments at a level that will lead to SUTVA violations
- You should consider compliance logistics: who needs to implement treatment assignments?
    - Teachers, doctors, tax collectors, etc. have different objectives and limited attention
- Are there political or social reasons to cluster treatment assignments?

For statistical power reasons, you want to randomize at the lowest level that works

Stratification

The Expected Level of Imbalance

Thought experiment:
You randomly assign treatment in a large sample, and then test 100 variables to see if they treatment and control group means are different – how many variables will be imbalanced?

Some imbalances (between treatment and control) matter more than others

- To enforce balance on important covariates, we typically **stratify** treatment assignments
- Intuitively, stratification is like running separate RCTs within each stratum
- In practice, we first sort by the stratification variables, then assign treatment

Stratification in Practice

First, choose what (dichotomous/categorical/binned) variable(s) to you want to stratify on

If you are using multiple stratification variables, number your strata in a reproducible way

The number of observations per strata must be at least as large as the number of treatments

Must have data on all stratification variables for all units being assigned to treatment

Reasons to Stratify

1. To enforce balance in terms of important covariates (e.g. baseline outcomes)
2. To enforce balance for fairness (or other political/feasibility) reasons
3. To enable tests of treatment effect heterogeneity
4. To increase statistical power (by explain variation in outcomes of interest)
    - You need to know what predicts your outcomes in cross-sectional data

Random Assignment in Practice: Questions and Takeaways

In the research design stage, ask yourself:

- What are you randomizing, and at at what level?
- Do you have a list of eligible units, and, if not, how will you create one?
- What are your stratification variables, and do you have data on them for all units?

Once you've collected the data described above, randomization is easy:

1. Set the seed(!), and then assign each eligible unit a pseudo-random number
2. Sort by stratum, and then by random number within each stratum
3. Assign treatments by counting off (0/1 or 1/2/3/etc.)
4. Check for balance, and make a balance check table