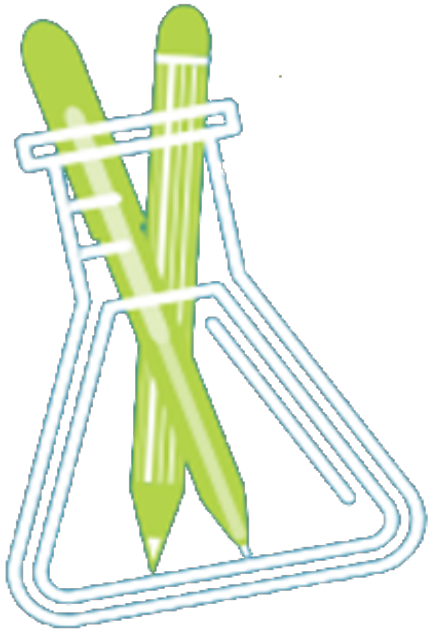
Measuring and Evaluating Determinants of Public Administration Productivity

Bureaucracy Lab

Development Impact Evaluation | Global Governance Practice

October 22-25, 2019, Brussels, Belgium



European Commission

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

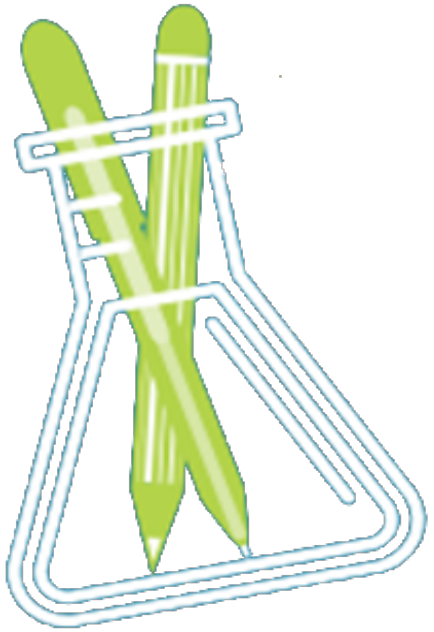
Measuring Impact III: Practical Considerations

Gianmarco León-Ciliotta
U. Pompeu Fabra, Barcelona GSE & IPEG

Bureaucracy Lab

Development Impact Evaluation | Global Governance Practice

October 22-25, 2019, Brussels, Belgium



European Commission

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

Outline

1. Defining your target population
2. Sample size and Power Calculations
3. Sampling methods
4. Timeline expectations
5. Budget calculations

Identifying your target population

- Who should we target in our data collection?
- The main source of guidance to define the target population will be the theory of change

1. Who is your treatment affecting?
2. What are the final outcomes we care about?
3. Which agents are involved in the causal chain of events?

Identifying your target population

Example:

Our intervention provides performance based incentives for the heads of primary health service facilities

- Target population of the intervention
- Outcomes of interest
- Agents involved in the causal chain of events

Identifying your target population

Additional considerations:

- Can we use administrative data?
  - If yes, this is ideal, since it will save us a lot of money!

  → Important considerations on the selection of the population included in this administrative datasets
- We are not able to get objective information, can we trust self reported data?
  - E.g. One of our target outcomes is individual level effort.
  - In case we have doubts, we can build into our design mechanisms to cross check these self-reports
    - E.g. Collect information from the "other" side of the market, build in a system of monitors, use "mistery" clients

Sample Size: Power Calculations

Why is sample size important?

- You want to estimate the impact of a certain program on your target population
- With less data, you will only get a noisy signal of this impact → not very informative
  - Type I Error (error of inclusion or false positive)
  - Type II Error (error of exclusion or false negative)
- *Main trade-off*: We would like to have data on the whole population, but this is unaffordable! (at least for me …)

→ How big should my sample be?

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Size of the expected impact

Sample Size: Power Calculations

Who is taller?

Increasing the sample acts as a magnifying glass to improve precision

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Size of the expected impact

- "What is the smallest effect size that, if it were any smaller, the intervention would not be worth the effort?"
- Large differences between groups are always easier to spot tan smaller differences!
- Q: How do I know whether I should expect a large or small impact? Moreover, what is large in this context?

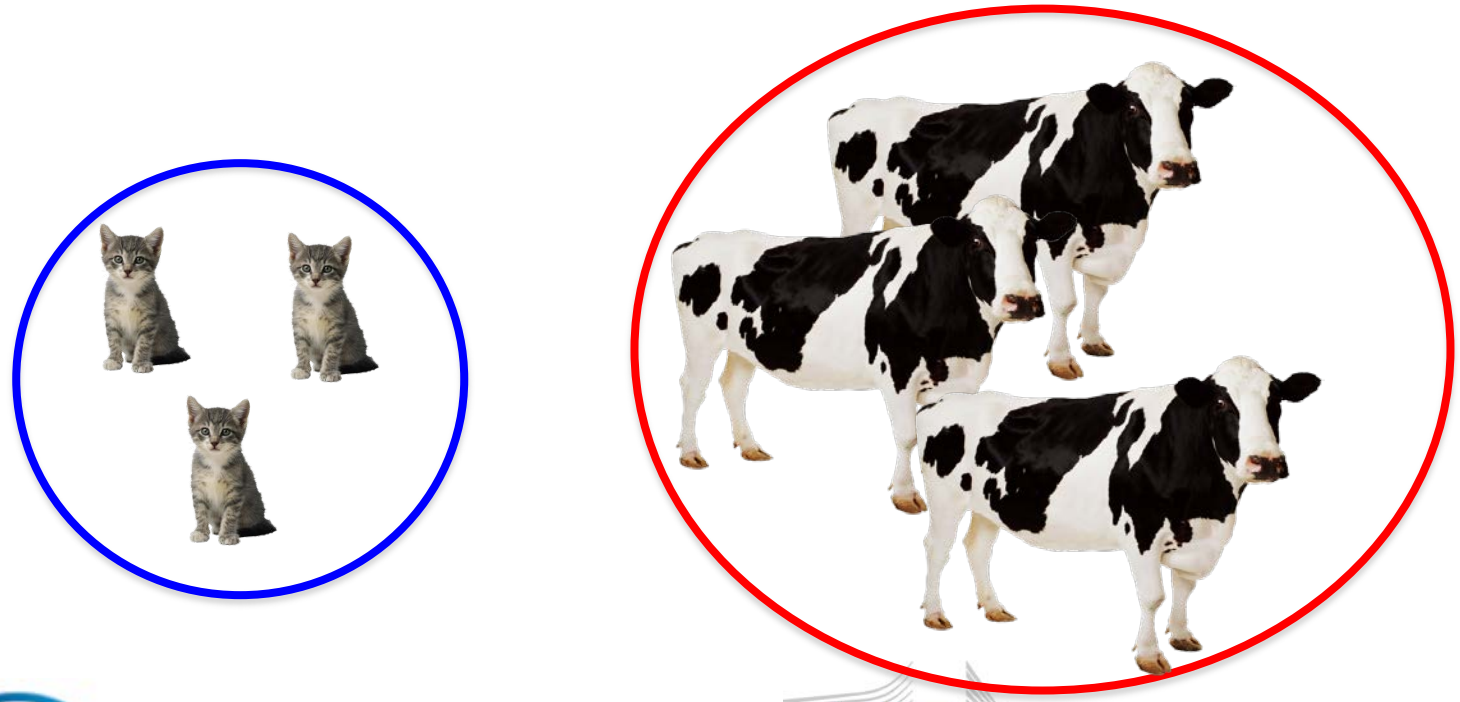
Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Variance of the outcome

Sample Size: Power Calculations

- Of the two (circled) populations, which animals are bigger?
- How many observations from each would you need to decide?



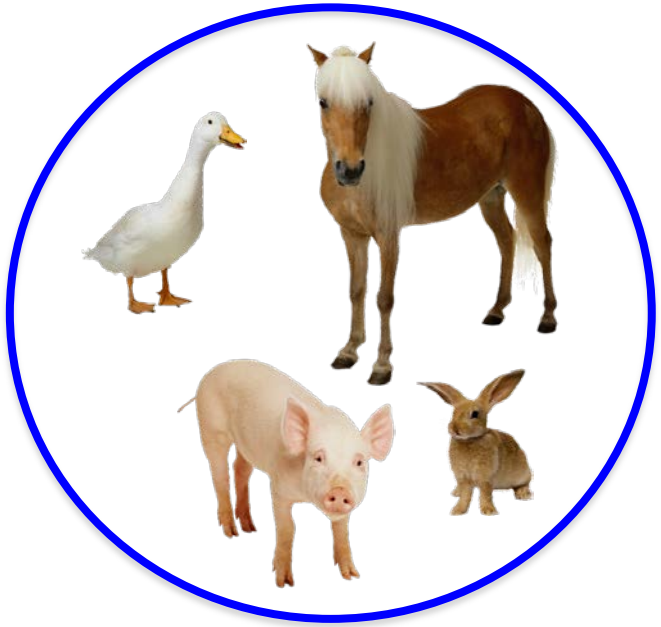
i2i DIME TRANSFORM DEVELOPMENT

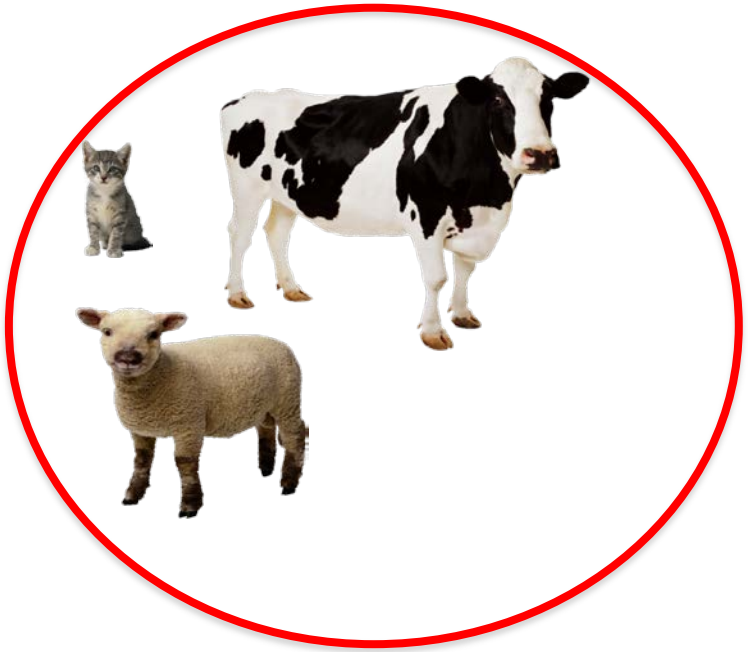
WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

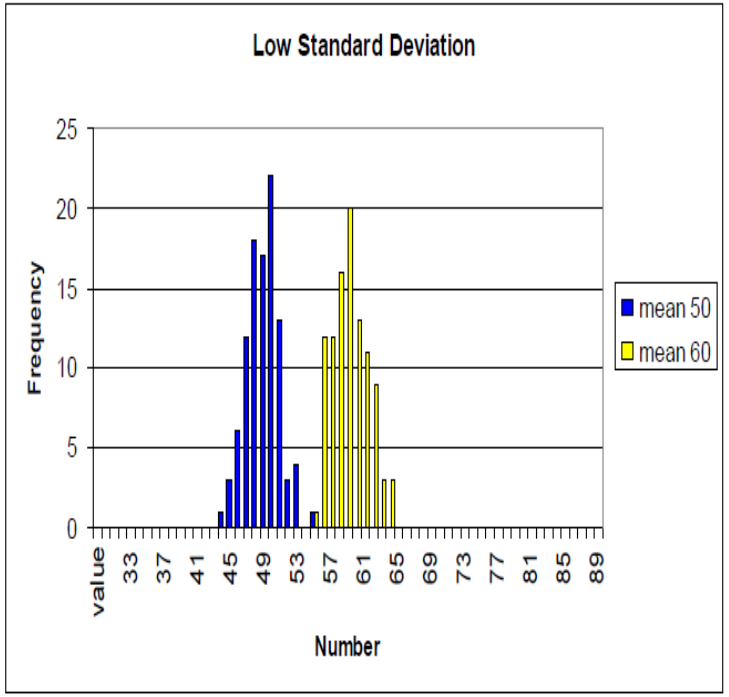
Does that change now? Why is it harder?

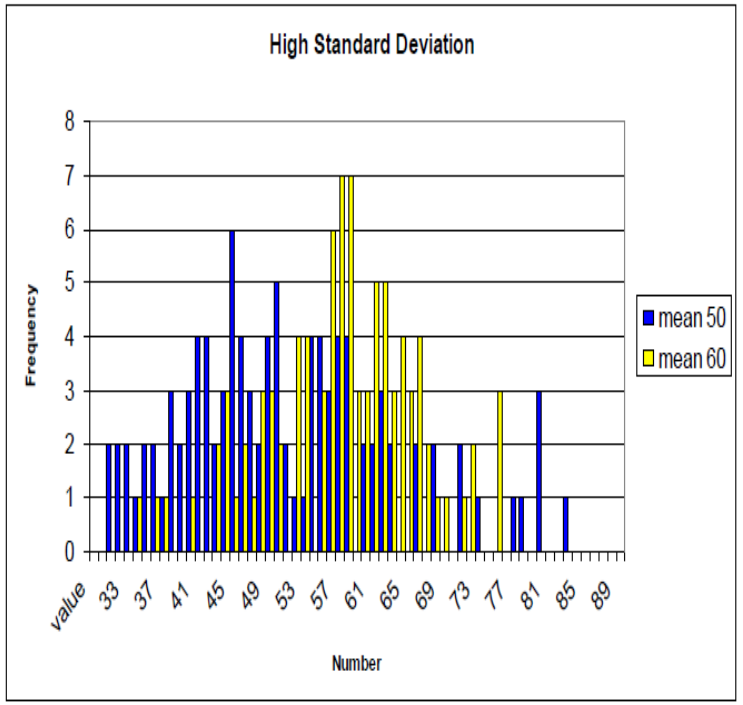




Sample Size: Power Calculations

In which case do we need a larger sample?





i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Variance of the outcome

- More underlying variance (heterogeneity) means that it is more difficult to detect difference and hence we need larger sample size

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Clustering

- Unit for sample size calculation depends on both:
  - Level of intervention AND
  - Level of measured impacts
- *Example*: intervention at clinic level, interested in impacts at HH level
  - Randomly assign clinics to treatment / control
  - Sample household within catchment areas
- Level of intervention ("cluster") most important for sample size calculation
- If few clusters, precision will be limited, regardless of number of HHs sampled

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

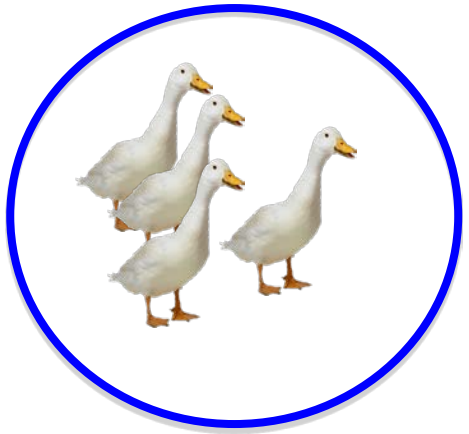
European Commission

Sample Size: Power Calculations

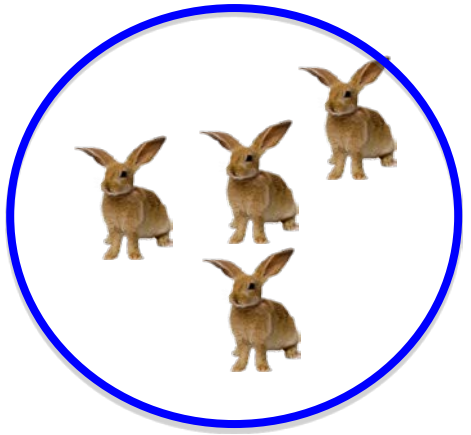
- Intracluster correlation (ICC): similarity of units within clusters
- Is the variation in outcome of interest coming mostly from differences within catchment areas (low ICC), or between catchment areas (high ICC)?
  - If HHs in catchment area 1 are similar to each other, but different from HHs in catchment area 2, high ICC
  - If HHs in catchment area 1 are similar to HHs in catchment area 2, low ICC

Sample Size: Power Calculations

Clustering (high ICC)



Area 1



Area 3

Area 2

Area 4

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

Clustering (low ICC)

Area 1

Area 3

Area 2

Area 4

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

Clustering:

- Larger within cluster correlation (guys in same cluster are similar) lower marginal value per extra sampled unit in the cluster

→ higher sample size/more clusters needed than a simple random simple

- Rule of thumb: at least 40 clusters per treatment arm (but always do your Power calculations first!)

Sample Size: Power Calculations

Question:

For an intervention randomized at the clinic (catchment area) level level, which sample design will have the most precision (highest power)?

A. 100 households from each of 2 catchment area
B. 25 households in each of 8 catchment area
C. 5 households in each of 40 catchment area

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Take-up

- Low take-up (rate) for intervention lowers precision
- Effectively decreases sample size / increases minimum detectable effect
  - Can only detect an effect if it is really large
- Unfortunately, to account for take-up rate of 50%, have to increase sample size by factor of 4

Sample Size: Power Calculations

What determines whether we want to have a larger or smaller sample?

Data Quality

- Poor data quality effectively increases required sample size
  - Missing observations
    - ✓quality of data collection, attrition, migration
  - High measurement error: answers not always precise
    - ✓e.g. self-reported land size, agricultural production
    - ✓e.g. recall bias, framing, pleasing
- Poor data quality can be partly addressed with field coordinator on the ground monitoring data collection

Sample Size: Power Calculations

Recap -- Things we should keep in mind:

- What is the …
  - level of randomization (clustering)?
  - Expected effect size?
  - Variation within target population?

- How to ensure …
  - High take-up?
  - Good data quality?

Sample Size: Power Calculations

Power calculations can be used in three ways :

1. to compute sample size, given power and minimum detectable effect size (MDES)
2. to compute power, given sample size and MDES, or
3. to compute MDES, given power and sample size.

Sample Size: Power Calculations

To calculate sample size with power calculations, you need to know

1. the MDES,
2. the standard deviation of the population outcome, and
3. the Type I and Type II significance levels.
4. For clustered samples, you need to also know the intra-cluster correlation and the average cluster size.

- *Key trade off to consider*: marginal value of added observations, i.e. accuracy vs. cost.

→ For detailed instructions on implementing power calculations, see Power Calculations in Stata and, as a compliment, Power Calculations in Optimal Design

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP Equitable Growth, Finance & Institutions

European Commission

Sample Size: Power Calculations

Data for Power Calculations and Key parameters

- You will likely never have all the data you need for your exact population of interest when you first compute power calculations
- You will need to use the best available data to estimate values for each parameter.
  - For MDE: Review the previous literature (relevant for the setting); trust knowledgeable actors in the field; minimum effect size for policy relevance
  - Ideal: pre-existing data … but often non-existent
  - Can use pre-existing data from a *similar* population
    - Example: LSMS, data routinely collected by govt, satellite imagery
  - Common sense

Sampling Methods

- Great! Now I have a simple size!

  …… Now, where do I get these people?

- *Best case scenario*: complete sampling frame already exists
- Most often this is not the case, especially since impact evaluation samples are quite specific
- First step is usually to conduct a listing, then sample
- However, that may not be entirely straightforward

Sampling Methods

Selected sampling methods:

Allow the enumerators to select the units to interview

- Pros:
  - Easiest for the enumerator
- Cons:
  - Enumerators are likely to chose the traders they can find easily
  - No guarantee of representation of all types of agents

Provide a walking pattern based on some criteria of coin flip/cell phone chosen, e.g. every X households knock on the door

- Pros:
  - There is some form of randomness
  - All types are likely to be represented
- Cons:
  - Enumerators have to do a lot of mental math!
  - Very hard to verify whether sampling pattern was followed

Sampling Methods

Selected sampling methods:

Rely on technology - Program the survey form to dynamically pick the units to survey

- Pros
  - Enumerators just have to locate the unit listed on the tablet screen
  - All types are likely to be represented
- Cons:
  - Programming of the randomization might take time
  - Randomization is not replicable if done on SurveyCTO

In case of an RD, sample observations around the cut-off

- If special RD, use satellite images, and send people to verify and run the surveys
- If based on a administrative threshold, sample observations around the cut-off using administrative data and interview them (power calculations needed)

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP
Equitable Growth, Finance & Institutions

European Commission

Budget calculations

What do I need to do?

- Step 1: Make a list of things your budget should include:
  - Salaries
  - Allowances
  - Transport and accommodation
  - Equipment
  - Stationery
  - Other

→ Make a list of all the things that will go into your budget

Step 2: Talk to people who have implemented surveys in your setting!

- How much do they pay for standard survey cost items?
- How do they organize transport for enumerators? Accommodation?

→Add item costs to the list you made

Budget calculations

- Step 3: Assumptions – Think through how fieldwork might be organized:
  - # surveys/person/day?
  - # teams I can realistically monitor?
  - time constraints
  - training duration
  - transport: car hire/public transport?
- Step 4: Bring it all together
  - Link each budget line item with the assumptions and with the standard rates
  - Add a buffer of extra survey days in case of delays. 15-20% extra is a good idea
  - Budget for contingencies: re-training, fuel price hikes, elections/political instability

→ Try out different scenarios to minimize costs!

i2i DIME TRANSFORM DEVELOPMENT

WORLD BANK GROUP Equitable Growth, Finance & Institutions

European Commission