Impact Evaluation Methods Experimental Design

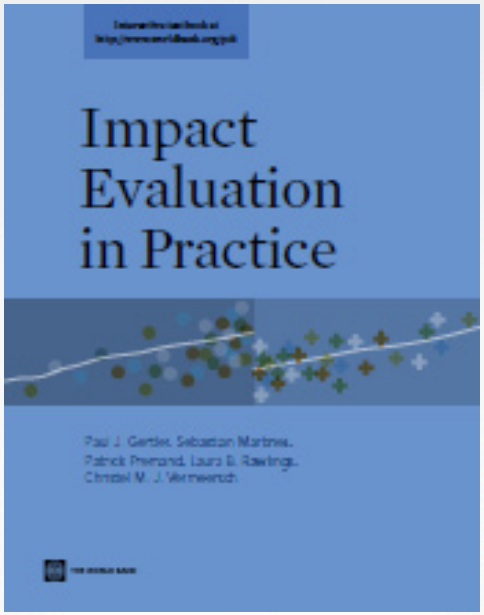
Presented by: Sergio Bautista-Arredondo
Instituto Nacional de Salud Pública

Enhancing Implementation Science
Program Planning, Scale-up, and Evaluation

Istituto Superiore di Sanità
Rome, Italy
July 16, 2011

HUMAN DEVELOPMENT NETWORK

SIEF



This material constitutes supporting material for the "Impact Evaluation in Practice" book. This additional material is made freely but please acknowledge its use as follows: **Gertler, P. J.; Martinez, S., Premand, P., Rawlings, L. B. and Christel M. J. Vermeersch, 2010, Impact Evaluation in Practice: Ancillary Material, The World Bank, Washington DC** (www.worldbank.org/ieinpractice). The content of this presentation reflects the views of the authors and not necessarily those of the World Bank.

Choosing your IE method(s)

Key information you will need for identifying the right method for your program:

Prospective/Retrospective Evaluation?

Eligibility rules and criteria?

- Poverty targeting?
- Geographic targeting?

Roll-out plan (pipeline)?

Is the number of eligible units larger than available resources at a given point in time?

- Budget and capacity constraints?
- Excess demand for program?
- Etc.

Choosing your IE method(s)

Choose the **best possible design** given the operational context:

Best Design

- Best comparison group you can find **+** least operational risk

Have we controlled for everything?

- Internal validity
- Good comparison group

Is the result valid for *everyone*?

- External validity
- Local versus global treatment effect
- Evaluation results apply to population we're interested in

Randomized Treatments & Comparison

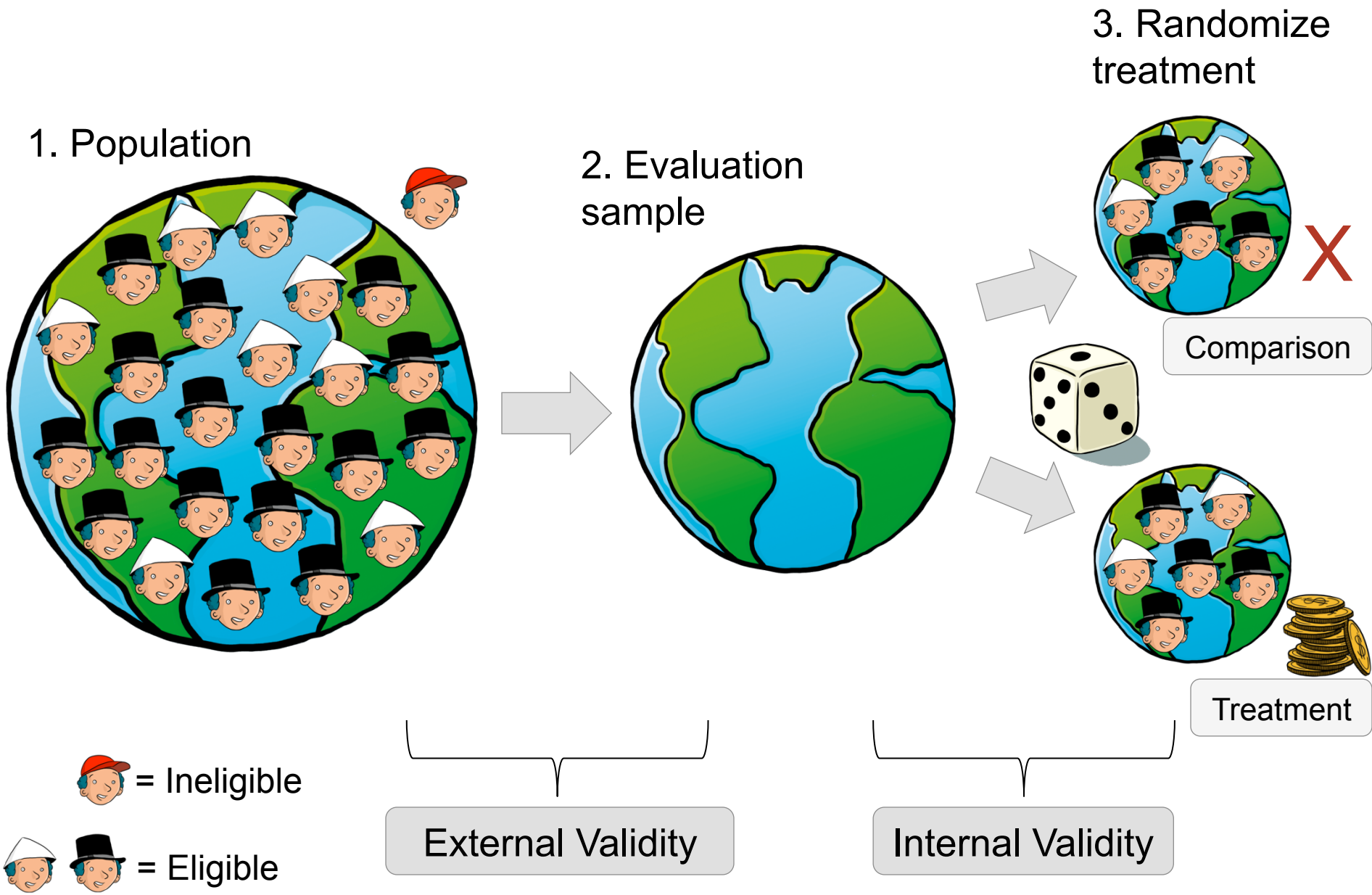
Eligibles > Number of Benefits Oversubscription

- Randomize.
- Lottery for who is offered benefits.
- Fair, transparent and ethical way to assign benefits to equally deserving populations.
- Give each eligible unit the same chance of receiving treatment.
- Compare those offered treatment with those not offered treatment *(comparisons).*

Randomized Phase In

- Give each eligible unit the same chance of receiving treatment first, second, third…
- Compare those offered treatment first, with those offered later *(comparisons).*

Randomized treatments and comparisons



Unit of Randomization

- Choose according to type of program
  - Individual/Household
  - School/Health Clinic/ catchment area
  - Block/Village/Community
  - Ward/District/Region

As a rule of thumb, randomize at the smallest viable unit of implementation.

- Keep in mind
  - Need “sufficiently large” number of units to detect minimum desired impact: Power.
  - Spillovers/contamination
  - Operational and survey costs

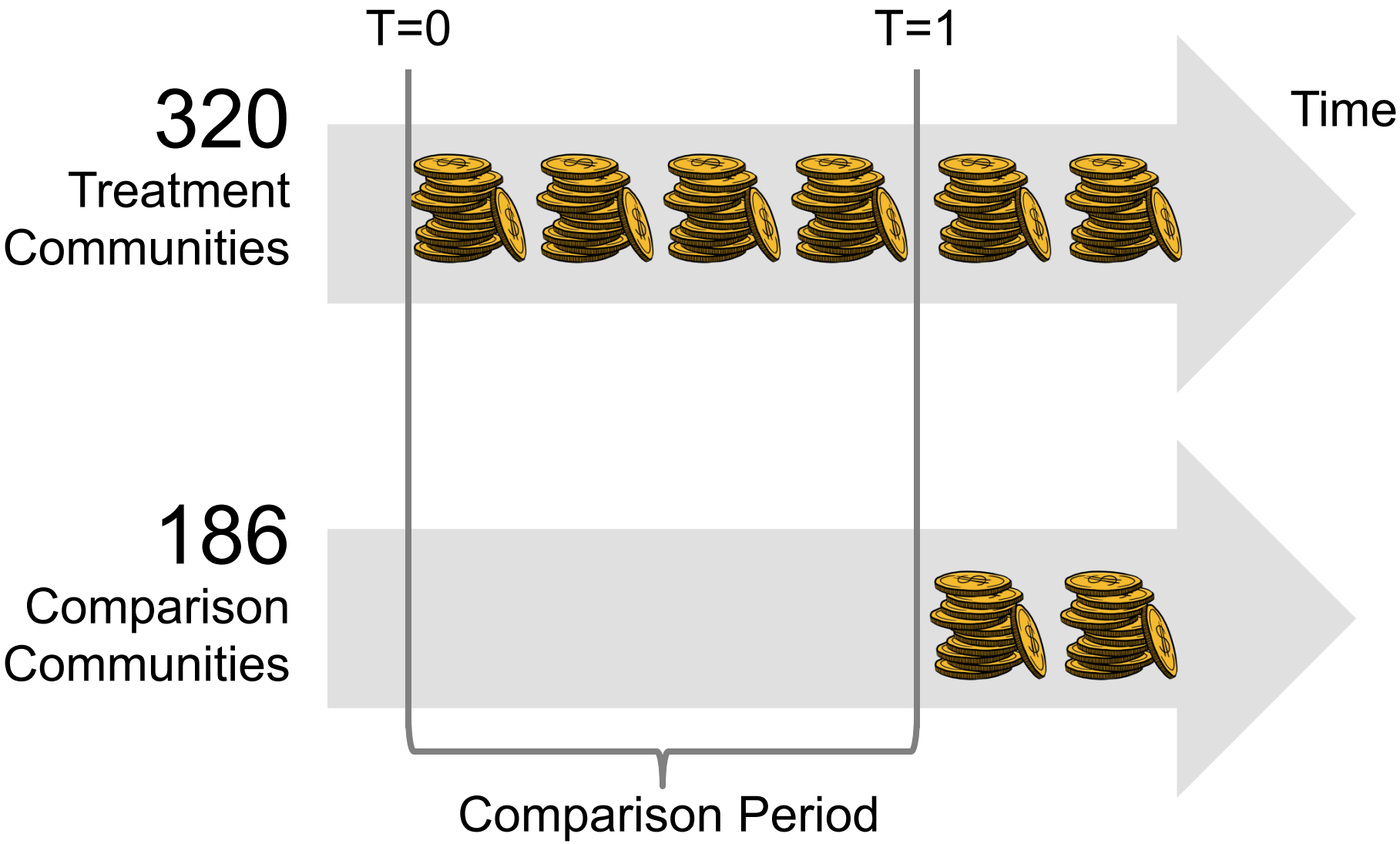
Randomized Assignment

- Progresa CCT program
- Unit of randomization: Community
- 506 communities in the evaluation sample
- Randomized phase-in
  - 320 treatment communities (14446 households):

    First transfers in April 1998.
  - 186 comparison communities (9630 households):

    First transfers November 1999

Randomized Assignment



Randomized Assignment

How do we know we have good clones?

In the absence of Progresa, treatment and comparisons should be identical

Let's compare their characteristics at baseline (T=0)

Balance at Baseline

Randomized Assignment			
	Treatment	Comparison	*T-stat*
Consumption ($ monthly per capita)	233.4	233.47	-0.39
Head's age (years)	41.6	42.3	-1.2
Spouse's age (years)	36.8	36.8	-0.38
Head's education (years)	**2.9**	**2.8**	**2.16****
Spouse's education (years)	2.7	2.6	0.006

Note: If the effect is statistically significant at the 1% significance level, we label the estimated impact with 2 stars (**).

Balance at Baseline

Randomized Assignment			
	Treatment	Comparison	*T-stat*
Head is female=1	0.07	0.07	-0.66
Indigenous=1	0.42	0.42	-0.21
Number of household members	5.7	5.7	1.21
Bathroom=1	0.57	0.56	1.04
Hectares of Land	1.67	1.71	-1.35
Distance to Hospital *(km)*	109	106	1.02

Note: If the effect is statistically significant at the 1% significance level, we label the estimated impact with 2 stars (**).

Randomized Assignment

	Treatment Group *(Randomized to treatment)*	**Counterfactual** *(Randomized to Comparison)*	**Impact** *(Y | P=1) - (Y | P=0)*
Baseline ***(T=0)*** Consumption (Y)	233.47	233.40	0.07
Follow-up ***(T=1)*** Consumption (Y)	268.75	239.5	29.25**

Estimated Impact on Consumption (Y)	
Linear Regression	29.25**
Multivariate Linear Regression	29.75**

Note: If the effect is statistically significant at the 1% significance level, we label the estimated impact with 2 stars (**).

Keep in Mind

Randomized Assignment

In Randomized Assignment, large enough samples, produces 2 statistically equivalent groups.

Feasible for prospective evaluations with over-subscription/excess demand.

We have identified the perfect **clone**.

Randomized beneficiary

Randomized comparison

Thank you!

sbautista@insp.mx