

# Evaluation of STDF-funded projects using Randomized Controlled Trials

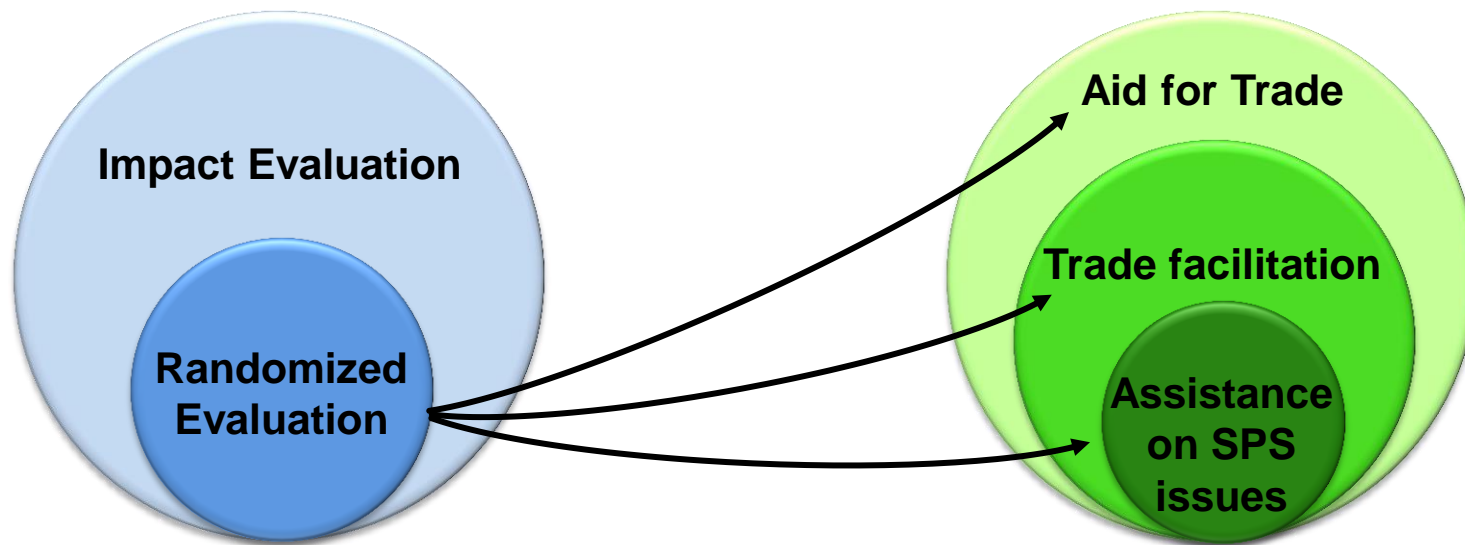
STDF Working Group Meeting  
World Trade Organization, March 28th, 2011

Vincent Pons (MIT), Frank Schilbach (Harvard)



Massachusetts Institute of Technology





# Road map of this presentation

1. Reasons for conducting such an impact evaluation
2. Broad questions the impact evaluation could answer
3. The randomized evaluation methodology: a simple way to produce undisputable results
4. Implementing a randomized evaluation: the flexibility given, and the constraints imposed by randomization

# 1. Reasons for doing such an impact evaluation

- Measure and understand the efficiency and cost-effectiveness of a particular intervention
  - ✓ How efficient is the intervention at achieving its goals? (agricultural development, public health, gain of market access abroad, poverty alleviation)
  - ✓ Why is the intervention efficient or inefficient?
  - ✓ How does its impact relate to its cost?
- Inform the design and selection of future interventions
  - ✓ Selection by STDF of interventions to be funded
  - ✓ Beyond STDF: Dissemination of evaluation results through STDF's platforms, partners and donors
- Ensure long-standing financial commitment from donors to Aid for Trade and SPS interventions
  - ✓ Cost-effectiveness measures give precise estimate of the value generated by the money invested
  - ✓ Use results to scale-up the most cost-effective programs only

## 2. The broad questions the evaluation could answer

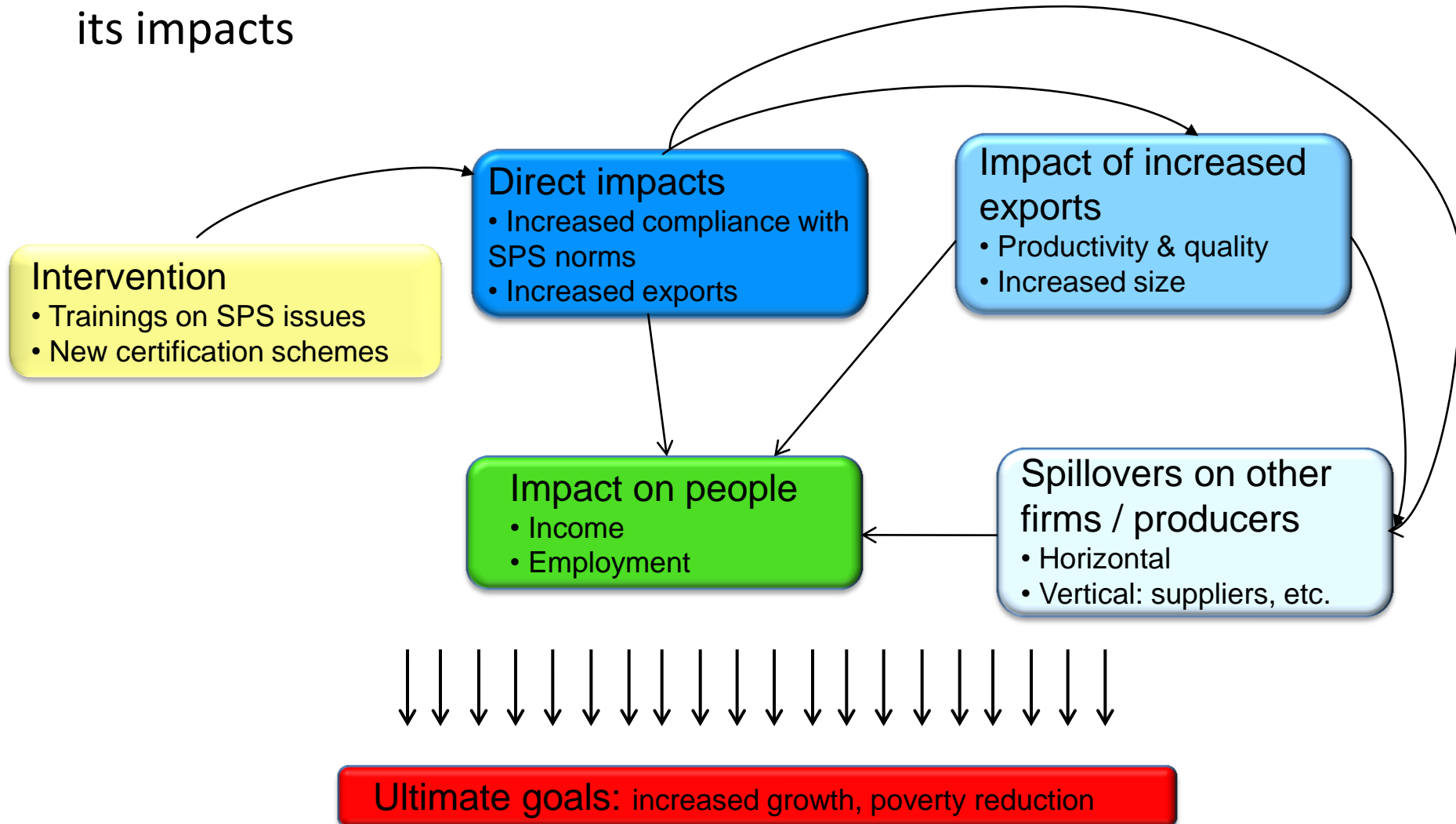
- Draw a complete causal chain between the intervention and all its impacts

The recent emphasis on Aid for Trade results from the belief that increased trade will promote growth and poverty reduction.

How does this translate at the micro level of a given project?

## 2. The broad questions the evaluation could answer

- Draw a complete causal chain between the intervention and all its impacts



## 2. The broad questions the evaluation should answer

- Measure the heterogeneity of the impact across several dimensions
  - ✓ The details of the intervention: training vs. training + follow-up
  - ✓ The types of producers / workers – in particular, does the intervention have an impact on the poorest; an identical impact among men and women?

# 3. The randomized evaluation methodology

- A methodology first used in clinical trials, now used extensively to measure the impact of education, health and other interventions in the developed and developing world
- The methodology in 4 steps
  1. Identify all farmers / cooperatives / firms eligible for the program and select a subsample of them for the evaluation (ex: 500 units)
  2. Randomly divide the subsample into 2 groups: Treatment (250 units) and Control (250). A coin is flipped for each unit.
  3. Administer the intervention to the Treatment group only.
  4. Measure and compare the outcomes between the Treatment and Control groups



# 3. The randomized evaluation methodology

## Evaluation without RCT

## With RCT

If the producers who benefit from the intervention are the largest / best informed

Outcome among beneficiaries –  
outcome among others  
= True impact + selection bias

Outcome in treatment – outcome in  
control (symmetric groups)  
= True impact

Trade lit: selection of the most  
productive / largest firms into  
exporting

If we expect long-term  
outcomes (e.g. impact  
of increased exports  
on productivity, size)

Difficult to control for time  
trends, other interventions,  
supply and demand shocks

Time trends and shocks are the  
same in T and C, thus controlled  
for

Trade lit: « treatment effect »  
of exporting

If we expect indirect  
outcomes (e.g. on the  
firm's employees, on  
other firms)

Difficult to control for  
confounding factors affecting  
these other layers

- The T and C's employees and suppliers are symmetric groups
- All shocks at all layers are controlled for

## 4. Implementing a randomized evaluation

- Make sure that the project is suitable for randomized evaluation:
  - ✓ Want more than 100 units, to be randomized between T and C
  - ✓ The intervention should not mechanically affect all units; some potential beneficiaries can be left out of the program, at least temporarily
    - Ex. of *unsuitable* project: new national or regional monitoring or animal and plant health surveillance system that affects all producers
  - ✓ Strong impact expected on limited number of identifiable beneficiaries vs. weak impact expected on large number of hardly identifiable beneficiaries
    - Ex. of *unsuitable* project: development of training material
    - Ex. of *suitable* projects: training programs (+ follow-ups) directly targeting local producers
- Get agreement of all stakeholders on the principle of randomizing
- Perform the random selection of beneficiaries EX ANTE

## 4. Implementing a randomized evaluation

- Randomization allows lots of flexibility
  - ✓ Ex. 1: the implementing organization wants a certain subsample of recipients to all receive the intervention immediately.
    - They will not be taken into account in the evaluation, but the randomization can be done on another subsample.
  - ✓ Ex. 2: the training program / new certification scheme should eventually be made available to all producers
    - Randomly select those who benefit from the program first (phase-in design)
    - Give more information / incentives to participate to the training to randomly selected producers (encouragement design)
    - Randomly allocate producers to different versions of the program (T1 vs. T2, not T vs. C)
- Cost of the evaluation
  - ✓ No extra-cost is imposed by randomization per se, compared to other impact evaluations.
  - ✓ The cost depends on the amount of data collected.