LECTURE 3: ALTERNATIVE STRATEGIES FOR RANDOMIZING PROGRAMS

World Agroforestry Center (ICRAF)
Nairobi, Kenya
January 23 – 26th, 2012

Overview

- How to measure impact?
  - The problem of the counterfactual
  - Experimental and non-experimental approaches

- We have decided we want to randomize, now how should we really do this?

How to measure impact?

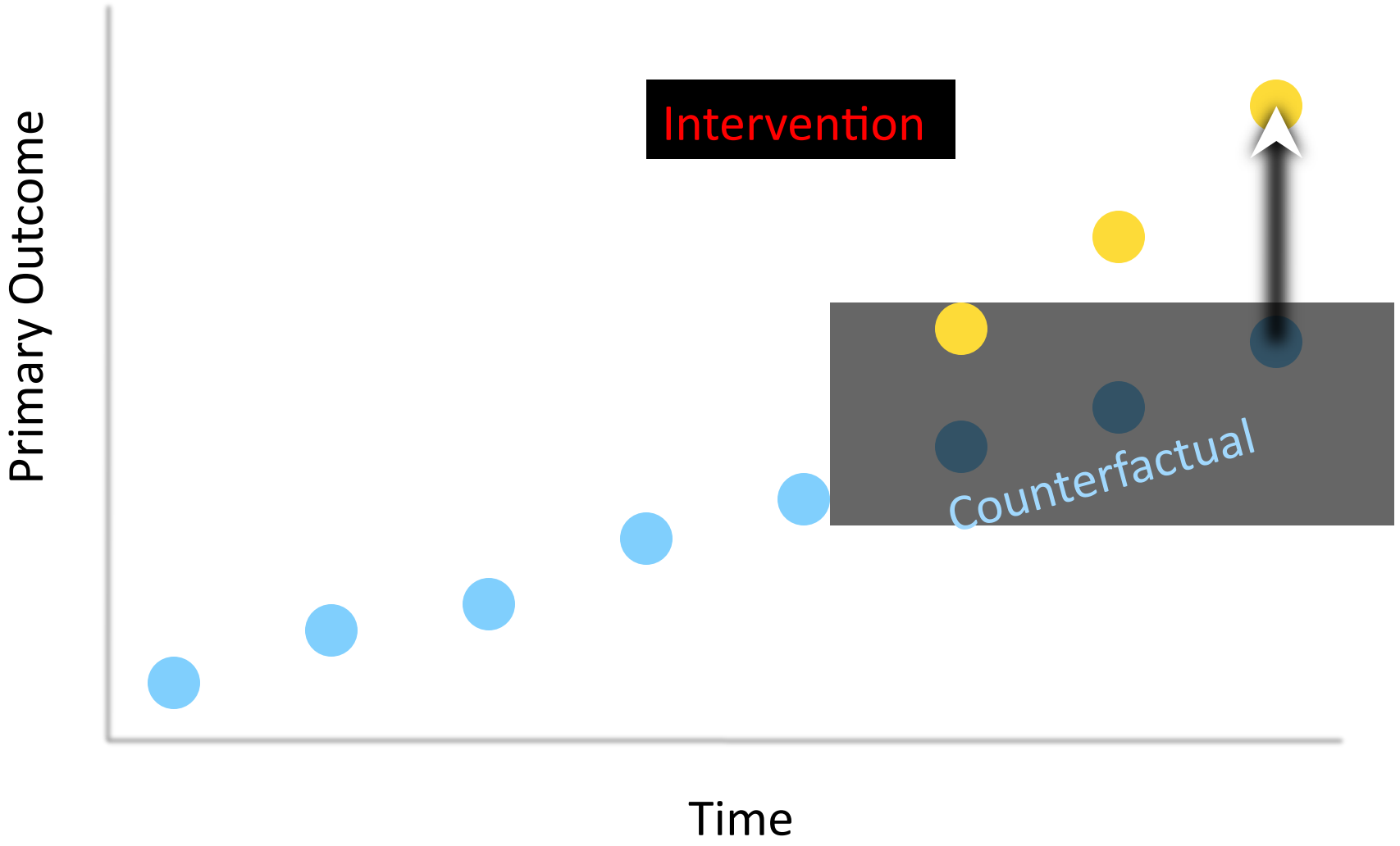
- What would have happened in the absence of the program?
- Take the difference between

  What happened (with the program) …and

  - *What would have happened (without the program)*

  = IMPACT of the program

Impact: What is it?



Assessing impact

- Examples
  - How much do extension services increase yields?
  - What are agricultural revenues with program providing market price information compared to without program?
- Compare same individual with & without programs at the same point in time
- BUT: Never observe same individual with and without program at same point in time

Solving the evaluation problem

- Counterfactual: what would have happened without the program
- Need to estimate counterfactual
  - i.e. find a control or comparison group
- Counterfactual Criteria
  - Treated & counterfactual groups have identical initial characteristics on average,
  - Only reason for the difference in outcomes is due to the intervention

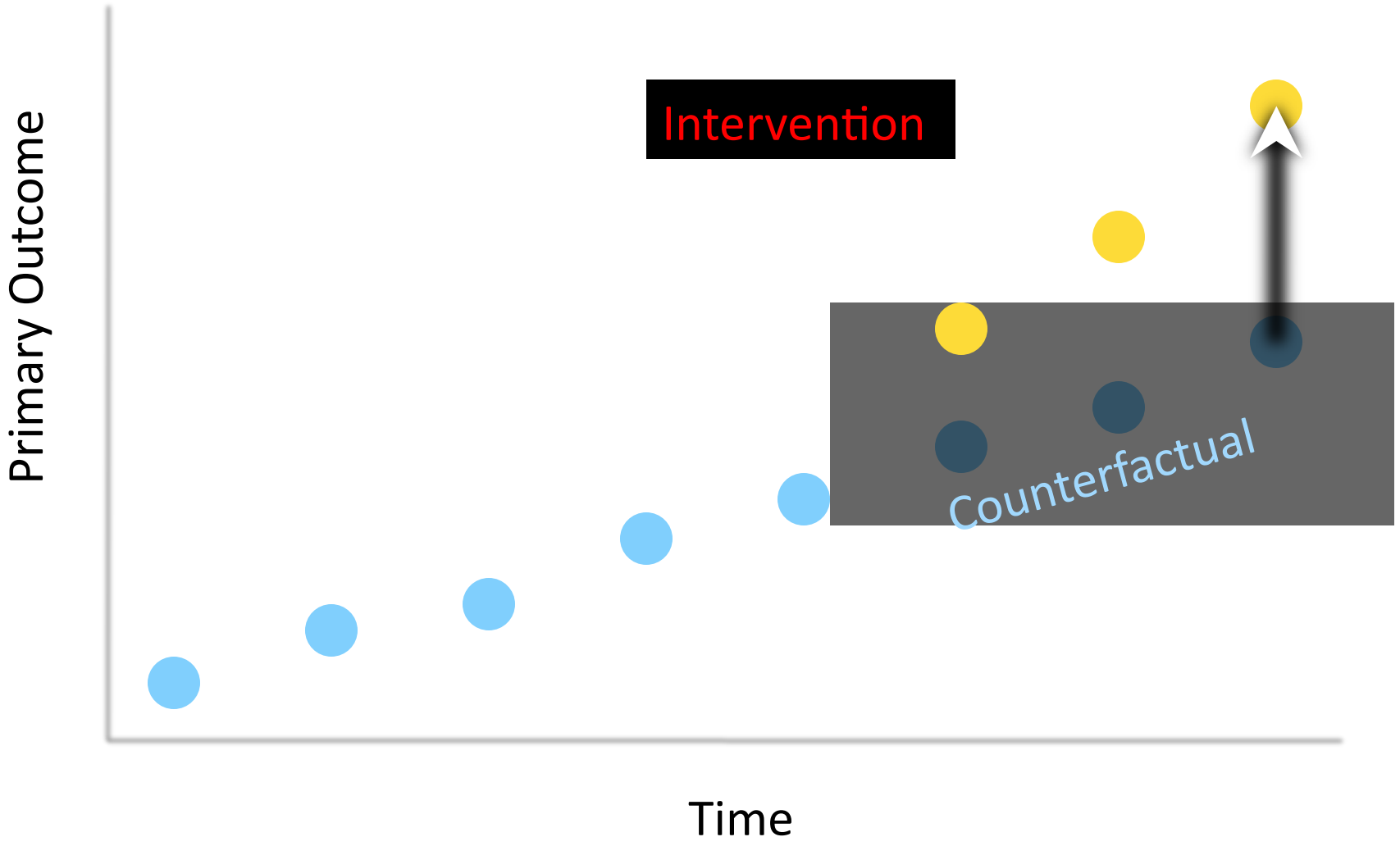
Constructing the counterfactual

- Counterfactual is often constructed by selecting a group not affected by the program
- Randomized:
  - Use random assignment of the program to create a control group which mimics the counterfactual.
- Non-randomized:
  - Argue that a certain excluded group mimics the counterfactual.

2 “candidate” counterfactuals

- Before and after:
  - Same individual before the treatment
- Non-Participants:
  - Those who choose not to enroll in program
  - Those who were not offered the program
- Question: why might neither of these be particularly good?

Before-after…?



Non-Participants….?

- Compare non-participants to participants
- Problem: why did they not participate?
- Selection bias
  - Self-selection
    - People choose to participate for specific reasons
  - Program placement
    - Programs target specific people, villages, areas for a reason
  - Many times reasons are directly related to the outcome of interest

  => Cannot separately identify impact of the program from these other factors/reasons

Possible Solutions…

- Guarantee comparability of treatment and control groups, so that
- ONLY remaining difference is intervention
- Different types of solutions
  - Experimental design/randomization/RCT
  - Quasi-experimental methods combined with econometric methods based on assumptions

Solutions hence involve…

- Randomization
  - Give all equal chance of being in control or treatment groups
  - Guarantees that all factors/characteristics will be on average equal between groups
  - Only difference is the intervention
- Or
  - need transparent & observable criteria for who is offered program
  - Econometric techniques with assumptions

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Randomization in practice: key steps

1. Design the study
2. Collect baseline data (?)
3. Randomly assign people to treatment or control
4. Verify that the assignment looks random (?)
5. Monitor process (~ is the randomization respected)
6. Collect follow-up data for T and C in identical ways
7. Estimate program impacts by comparing mean outcomes in T and C
8. Assess whether program impacts are statistically and practically significant

=> All the hard work happens up front

Prospective designs

- Use opportunities to generate good control groups
- Most programs cannot deliver benefits to all those eligible
  - Budgetary limitations:
    - Eligible who get it are potential treatments
    - Eligible who do not are potential controls
  - Logistical limitations:
    - Those who go first are potential treatments
    - Those who go later are potential controls

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Methods of randomization

- Basic lottery
  - Useful when program is oversubscribed; ok for some to get nothing
- Random Phase-in
  - Useful when everybody must get treatment eventually
- Random Rotation
  - Everybody must get something at some point, but not enough resources for all at the same time
- Encouragement
  - Program has to be open to all, but take up in general is low and can be impacted (a lot) with incentives

=> Advantages and disadvantages of each approach?

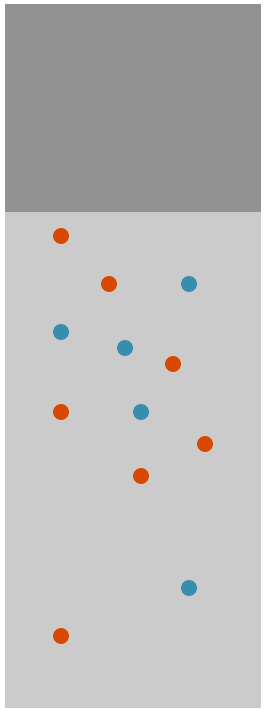
Examples in agriculture

- Basic Lottery
  - Lottery to receive fertilizer vouchers
- Random phase-in (everyone gets it eventually)
  - Train some farmers groups each year
- Random rotation
  - Some get free inputs in one year, others get it in second year, …
- Encouragement design
  - One farmers support center per district
  - Some farmers get travel voucher to attend the center

Will lotteries be “acceptable” ?

- Often (not always) perceived as fair and transparent
- Oversubscription can help (often many more eligible people than resources available)
- But also, we can make them more acceptable
  - Program selects those that they would always select
  - Others (threshold cases) are allocated via lottery
  - Questions about which group does this represent (~ external validity)

Lottery among the qualified



Must get the program whatever

Randomize who gets the program

Not suitable for the program

Phase-in: takes advantage of expansion

- Everyone gets program eventually
- Natural approach when expanding program faces resource constraints
- What determines which schools, branches, etc. will be covered in which year?

Phase-in designs

Advantages

- Everyone gets something eventually
- Provides incentives to maintain contact

Concerns

- Can complicate estimating long-run effects
- Care required with phase-in windows
- Do expectations of change actions today?

Rotation design

- Groups get treatment in turns
- Advantages
- Concerns

Encouragement design: What to do when you can't randomize access

- Sometimes it's practically or ethically impossible to randomize program access
- But most programs have less than 100% take-up
- Randomize encouragement to receive treatment

What is "encouragement"?

- Something that makes some folks more likely to use program than others
- Not itself a "treatment"
- For whom are we estimating the treatment effect?
- Think about who responds to encouragement

To summarize: Possible designs

- Simple lottery
  - Randomization in the “bubble”
- Randomized phase-in
- Rotation
- Encouragement design

  - Note: These are not mutually exclusive.

Method used in part depends on opportunities for randomization

- Budget constraint prevents full coverage
  - Random assignment (lottery) is fair and transparent
- Limited implementation capacity
  - Phase-in gives all the same chance to go first
- No evidence on which alternative is best
  - Random assignment to alternatives with equal *ex ante* chance of success

Method used in part depends on opportunities for randomization (2)

- Take up of existing program is not complete
  - Provide information or incentive for some to sign up
- Pilot a new program
  - Good opportunity to test design before scaling up
- Operational changes to ongoing programs
  - Good opportunity to test changes before scaling them up

Methods: advantages and disadvantages

- Basic lottery
  - Adv: easy to understand and implement; can be done in public
  - Disadv: control group might not cooperate; differential attrition
- Phase-in
  - Adv: easy to understand and explain (resource constraint) + control group likely participates
  - Disadv: anticipation effects and can't get long-term impacts
- Rotation
  - Adv: more observations than phase-in
  - Disadv: can't get long-term impacts
- Encouragement
  - Adv: can randomize at individual level
  - Disadv: impact of those that respond to incentives; and impact of incentives on take up has to be high. Encouragement can also have direct effect

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Further variations to learn even more

- Multiple treatments
- Cross-cutting treatments
- Varying levels of treatment

Multiple treatments

- Sometimes core question is deciding among different possible interventions
  - Might very well be that ex-ante it is unclear which treatment is going to work better
  - E.g. Information on new technology to men or to women?
  - E.g. credit or training?

  => You can randomize these programs

- Does this teach us about the benefit of any one intervention?
  - Do you have a control group?
- Advantage:
  - Can help to disentangle impacts of interventions that can be considered a "package" (~ CCTs)
  - Can shed light on complementarities and substitutabilities
  - Can help shed light on operational questions, go beyond simple "impact"

Cross-cutting treatments

- Test different components of treatment in different combinations
- Test whether components serve as substitutes or compliments
- What is most cost-effective combination?
- Advantage: win-win for operations, can help answer questions for them, beyond simple “impact”!

Varying levels of treatment

- Some schools are assigned full treatment
  - All kids get pills
- Some schools are assigned partial treatment
  - 50% are designated to get pills
- Testing subsidies and prices

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Different units you can randomize at

- Individual
- Farm
- Farmers' Association
- Irrigation block
- Village level
- Women's association
- Youth groups
- School level

Unit of randomization

- Randomizing at higher level sometimes necessary:
  - Political constraints on differential treatment within community
  - Practical constraints—confusing for one person to implement different versions
  - Spillover effects/externalities may require higher level randomization
    - Need to guarantee that individuals in the control are not affected by the treatment

- Randomizing at group level requires many groups because of within community correlation

How to Choose the Level

- Nature of the Treatment
  - How is the intervention administered?
  - What is the catchment area of each "unit of intervention"
  - How wide is the potential impact?
- Aggregation level of available data
- Power requirements
- Generally, best to randomize at the level at which the treatment is administered.

Level of randomization

- More generally: if there can be spillovers (positive or negative, need to randomize at higher level)
- Higher-level randomization might also be needed for reasons of acceptance/fairness
- BUT: statistical power of individual randomization is higher (bigger sample size at lower cost)
- One possibility: encouragement design or 2 levels of randomization
  - E.g. program versus non-program: community level
  - 2nd randomization within community: variation because of different variations in program, or different encouragements, …

    => allows for specifically studying spillovers

    (sometimes spillovers are also studies by considering non-eligible)

Further on spillovers

- What if spillovers to other communities?
- Can use random variation in density/proximity to treatment
  - Examples: de-worming
- Spillovers are not always obvious and can be hard to predict

  => more generally: "social" effects are hard to identify :

  do people that live close to each other behave similarly because

  1) they are likely similar to each

  2) they are exposed to the same shocks/exogenous factors

  3) they influence each other behavior

  ⇒ randomized experiments can shed light on them because one create exogenous variation at various levels using different levels of randomization

  ⇒ Important for social learning (e.g. technology adoption)

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Ethical Considerations

- Do not delay benefits: Rollout based on budget/ administrative constraints
- Equity: equally deserving beneficiaries deserve an equal chance of going first
- Transparent & accountable method
  - Give everyone eligible an equal chance
  - If rank based on some criteria, then criteria should be quantitative and public

Constraints: Politics

- Need to know political dynamics and try to anticipate !
  - Guarantee ex-ante buy-in into evaluation design by key decision makers

- Some factors can help
  - Lotteries are simple, common and transparent
  - Randomly chosen from applicant pool
  - Participants know the “winners” and “losers”
  - Simple lottery is useful when there is no a priori reason to discriminate
  - Transparent
  - Often perceived as fair

Constraints: Resources

- Most programs have limited resources
  - Vouchers, Farmer Training Programs
- Results in more eligible recipients than resources will allow services for
- Are often an evaluator's best friend
  - Key is to get around mechanism implementers use to match recipients to treatment
  - Recruit until constraints met
  - May distinguish recipients by arbitrary criteria

Constraints: contamination Spillovers/Crossovers

- Remember the counterfactual!
- If control group is different from the counterfactual, our results can be biased
- Can occur due to
  - Spillovers
  - Crossovers

Constraints: logistics

- Need to recognize logistical constraints in research designs.
- Ex: training in new technology with extension agents
  - Suppose this training is one of many responsibilities of extension agents
  - Suppose the extension agents served members from both treatment and control groups
  - It might be difficult to train them to follow different procedures for different groups, and to keep track of what to give whom

Constraints: sample size

- The program is only large enough to serve a handful of communities
- Primarily an issue of statistical power
- Will be addressed tomorrow

Outline – second part

- We have decided we want to randomize, now how should we really do this?
  - Steps in randomization
  - Methods of randomization
  - Variations on simple treatment-control
  - Levels of randomization
  - Real world concerns and constraints
  - Some examples

Some examples

1) Household risk management and adaptation to climate change in Nicaragua

2) Fertilizer adoption in Kenya

3) Improved seed adoption in DRC

Example 1:Households risk management adaptation to climate change

- "Atencion a Crisis" Program of the ministry of the family (MIFAMILIA)
- 6 municipalities in rural Nicaragua with high levels of extreme poverty and frequent droughts
- 3000 poor and vulnerable households
- Combine conditional cash transfers with interventions targeted at increasing the productive capacity of poor households
  - 1000 hh: CCT
  - 1000 hh: CCT + vocational training
  - 1000 hh: CCT + productive investment grant
- November 2005 - December 2006
- Randomized selection in two steps, at 2 different levels
  - Random Control (50) and Treatment communities (56)
  - Within treatment communities: Lottery to select families in each of the 3 packages

Design/Timing of the Impact Evaluation

- Baseline in 2005
  - No baseline differences between treated and control households, nor between different treatment groups
- Follow up survey – July-August 2006
  - 9 months after the program began
- Program ends December 2006
- Second follow-up survey in 2008-2009

  ~ 2 years after end program
- Third follow-up survey in 2011 (ongoing)

=> Allows for dynamic analysis of adoption/adaptation and analysis of sustainability of program impacts

Beneficiaries of the productive investment package

Main findings on risk management

- Beneficiaries are better protected against negative impact of weather shock
  - On total consumption and food consumption
  - Protection highest for productive transfers (T3)
- 2 years after the end of the intervention, the impact of productive transfers significantly positive, and significantly larger than other interventions in presence of shocks
- Mechanisms?
  - Program induced income diversification
    - More livestock activities/income
    - More nonagricultural activities/income
  - Program leads to changes in attitudes towards nonagricultural and agricultural activities

    => Role of leaders identified through community level randomized variation

Example 2. Fertilizer in Kenya

Farmers in Western Kenya could get 70% returns using fertilizer

Only 37 percent use it

Many have used it before and have seen the returns

They can buy small quantities if they don't have much money

What could be the problem?

- **Technology?**
- **Information**
- **Procrastination**

Experimental design

- Series of experiments on same population
- Multiple treatments in various years
  - Earlier years: demonstration plots of self and neighbor, as well as school demonstration plots

    => very little "learning from doing"

    => and little "learning from others"
  - Follow-up study: Multiple treatments
    - Variations of savings commitment device
      - SAFI offered at harvest
      - SAFI with ex-ante choice of timing, offered even before harvest
    - Comparison with other treatments to rule out certain mechanisms
      - Offered at time of application
      - Offered at time of application with discount

Example 3: Developing improved seed markets in DRC

- Multi-arm project aimed at improving agricultural productivity through road rehabilitation, re-introduction of improved seeds and animals, training, market and storage construction, etc…
- Some of these are too political, too few (markets), technically driven, etc… for an RCT
- In particular: one of the goals : developing the seed market in Equator (there is only 1!)

But…

- Let's think of final outcomes: agricultural productivity and household welfare
- If we want availability of improved seeds to lead to such final outcomes, we need:
  - Households to
    - adopt the improved seeds (i.e. buy them),
    - believe it is profitable for them (hence need information both about the returns to the seeds, but also commercialization possibilities),
    - Have improved seeds lead to higher yields and increased consumption
  - Might depend also on intrahousehold dynamics

All of a sudden many RCTs…

- Or more specifically, many plausible variations in the design of the specific program components that indicate:
  - Existing questions on economic behavior and therefore on optimal program design
  - Not all will happen at once
  - Not clear a priori which is the best way to go
  - ⇒ Answering those questions can help in improving design of subsequent phase of the interventions
  - ⇒ Can create a first experience with rigorous IE

Questions?