INTERNATIONAL FOOD POLICY
RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty
Supported by the CGIAR

EVALUATING THE IMPACT OF NEW AGRICULTURAL TECHNOLOGIES WITH PANEL DATA

JOHN HODDINOTT

DECEMBER 3, 2010

The Four Horseman of the Evaluation

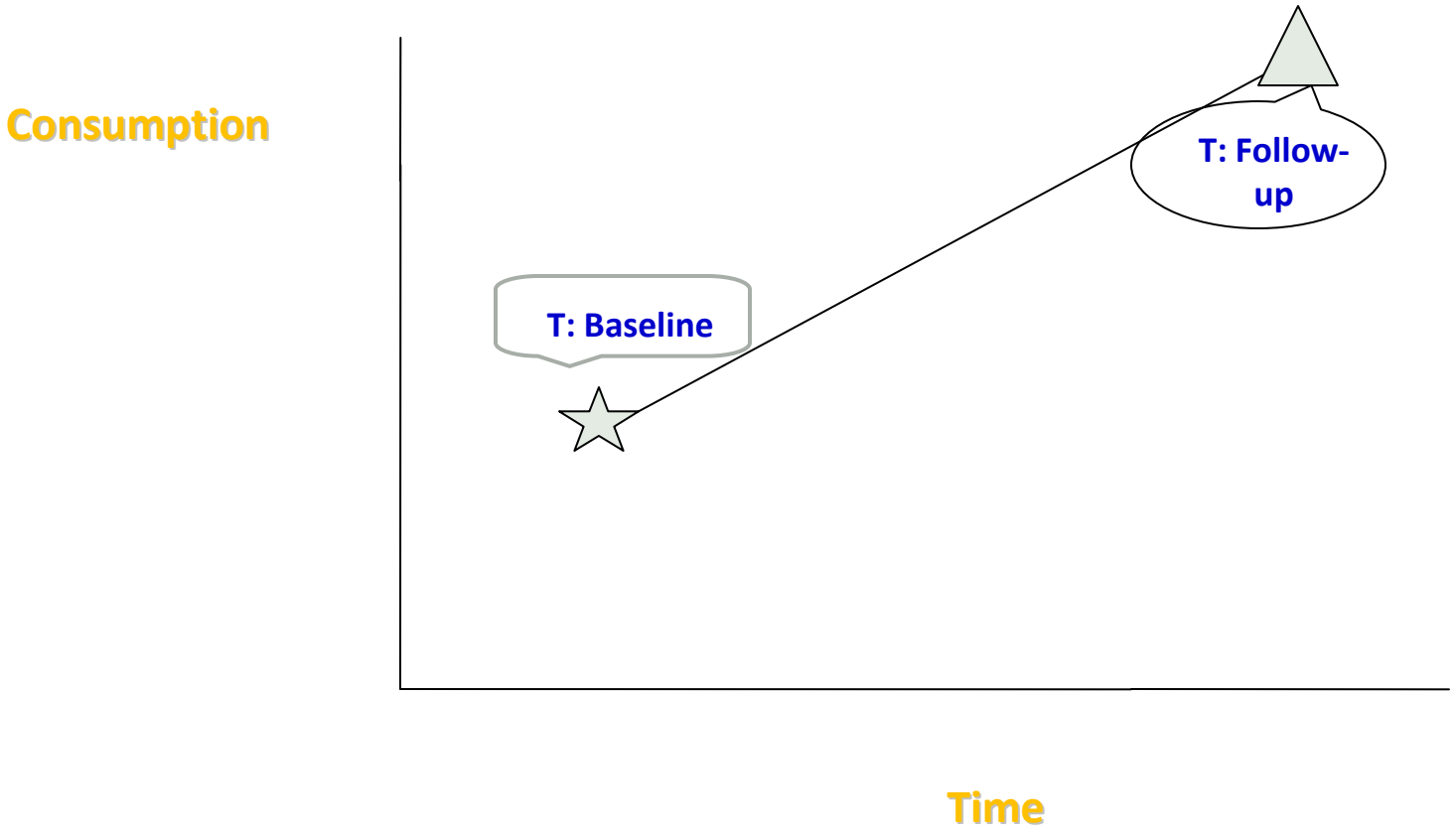
- The problem of the counterfactual:
  - What would have happened in the absence of the intervention

- "Suppliers" of the innovation may target potential beneficiaries on the basis of both characteristics that the evaluator can observe and those that she cannot observe

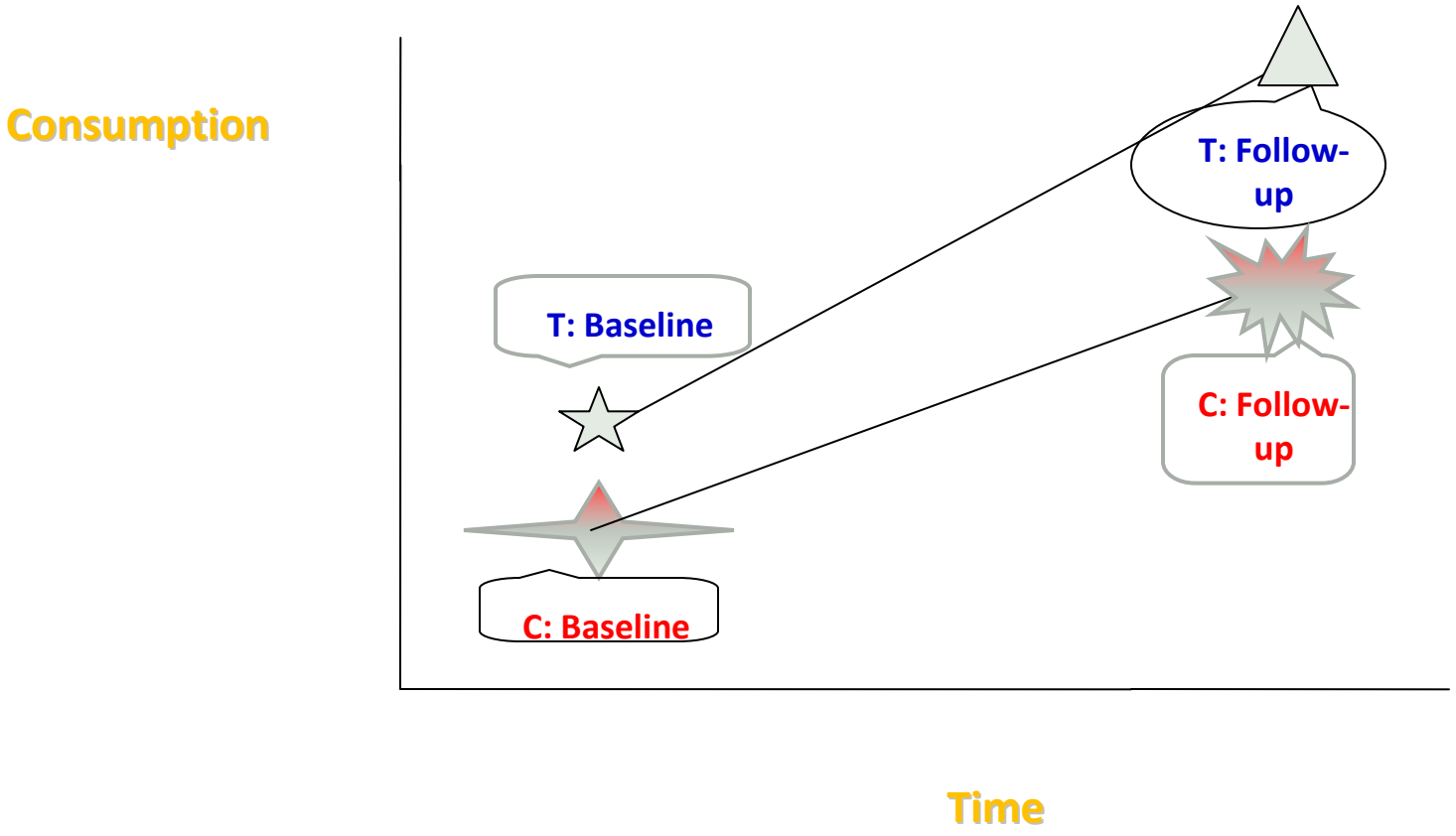
- "Adopters" will differ from non-adopters in terms of both observable and non-observable characteristics

- Adoption is not necessarily a binary action. There can be partial adoption:
  - Using the innovation on some, but not all, land
  - Using the innovation less intensively than that recommended by its designers (eg, using less than the recommended level of fertilizer)

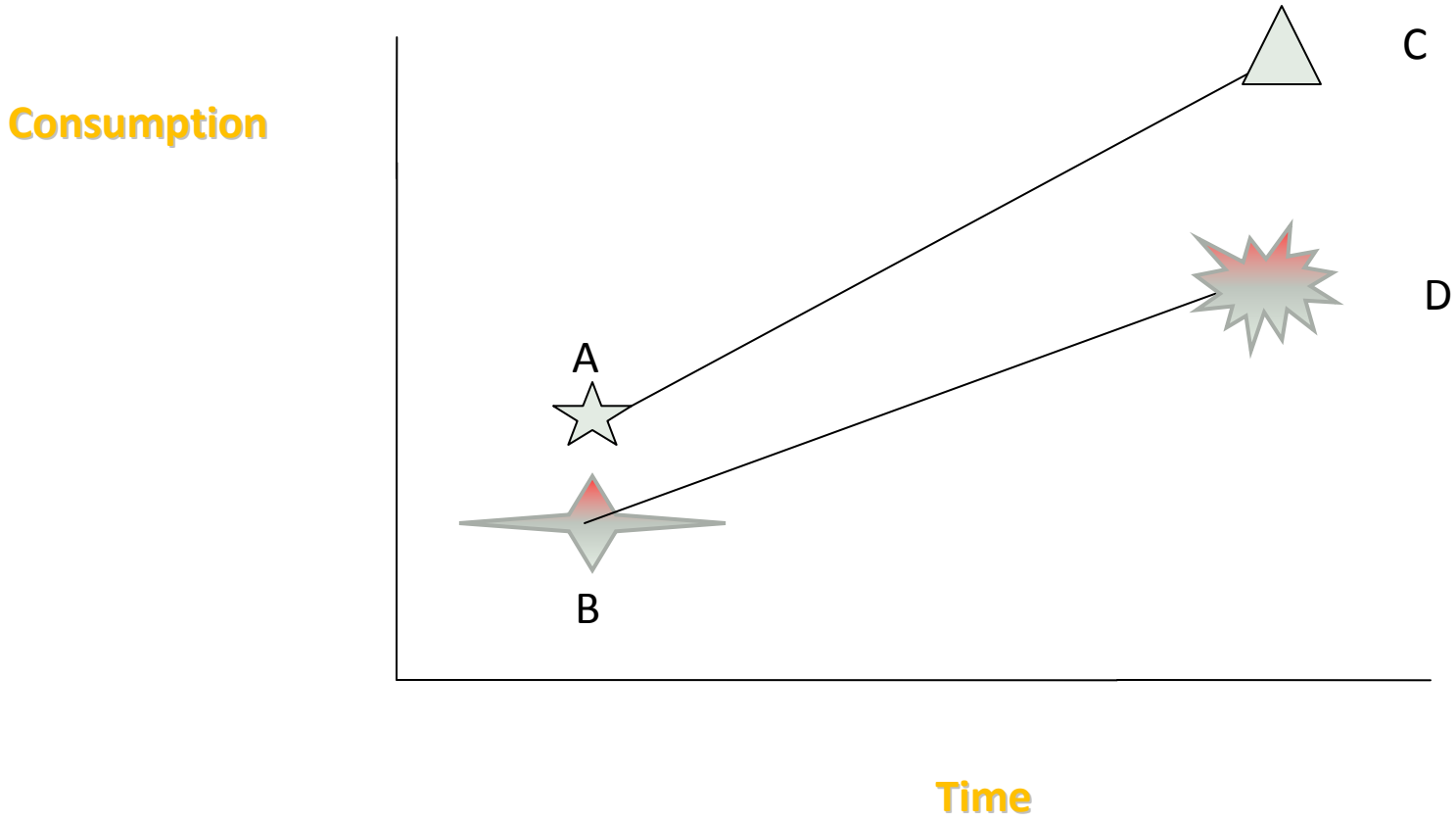
The Problem of the Counterfactual



The Problem of the Counterfactual



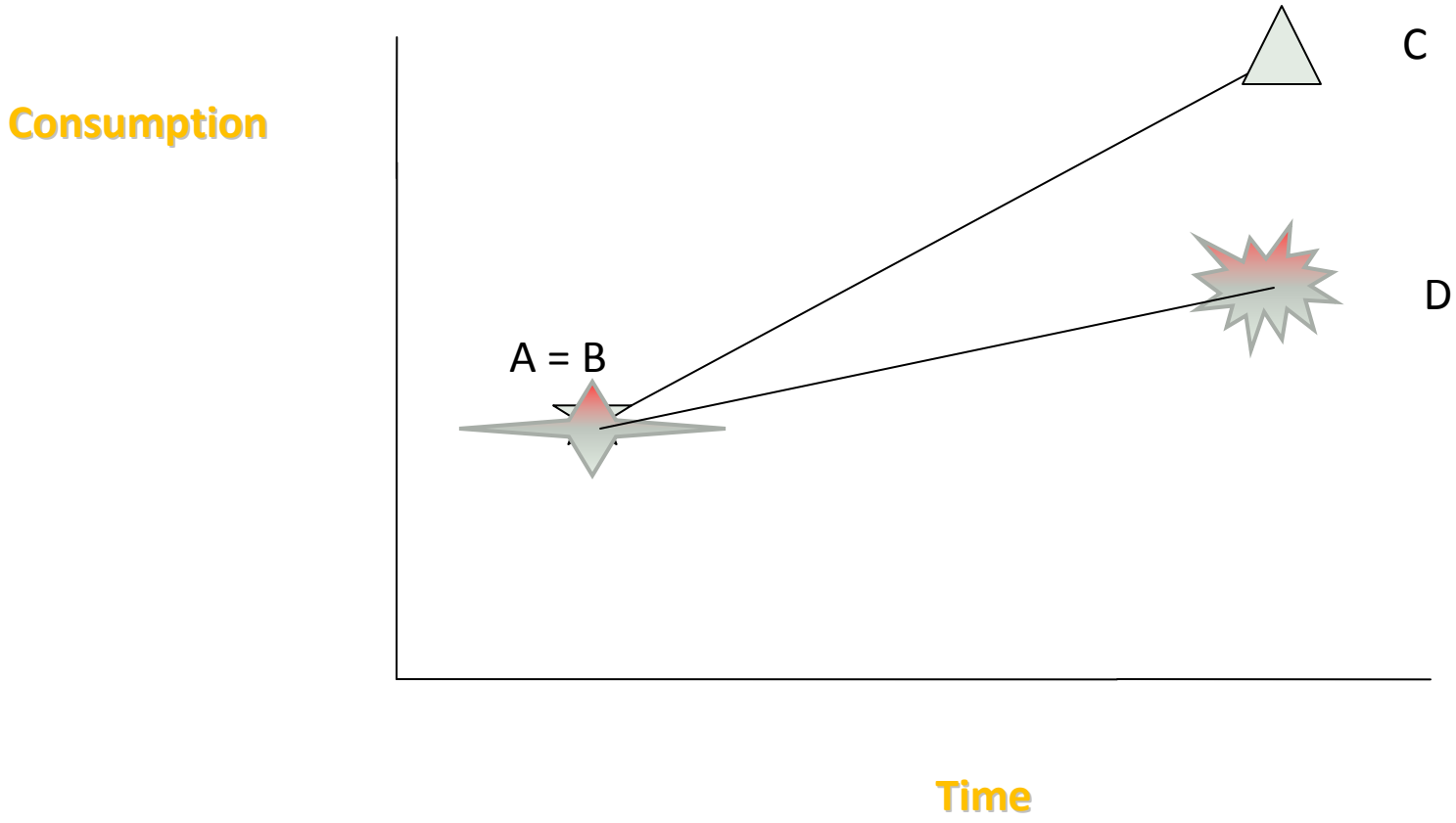
The Problem of the Counterfactual



"Ex post" impact = (C – D) "Single difference"

"True" Impact = (C – A) – (D – B) "Double Difference"

The Problem of the Counterfactual



If A=B, then Impact = (C – A) – (D – B)

= (C – D) "Single Difference"

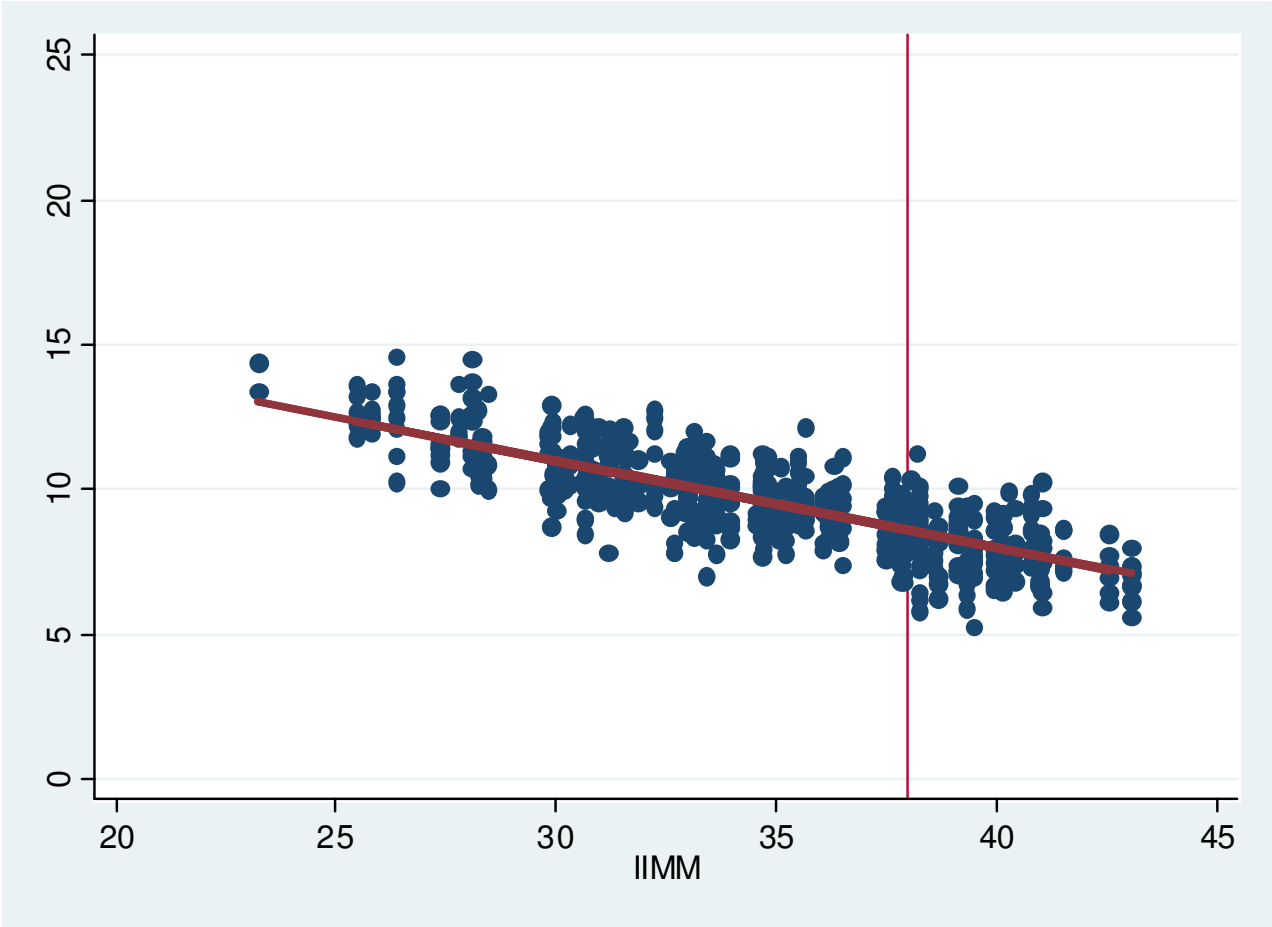
The four horseman, redux

- Panel data – using “before and after” as well as “with and without” – allows us to assess impact without having to make the assumption of equality of outcomes at baseline

- But the collection of panel data by itself does not address the other problems noted in the introduction. In the absence of experimental approaches such as RCTs (already discussed), they need to be combined with a non-experimental method:
  - Regression discontinuity methods.
  - Matching methods

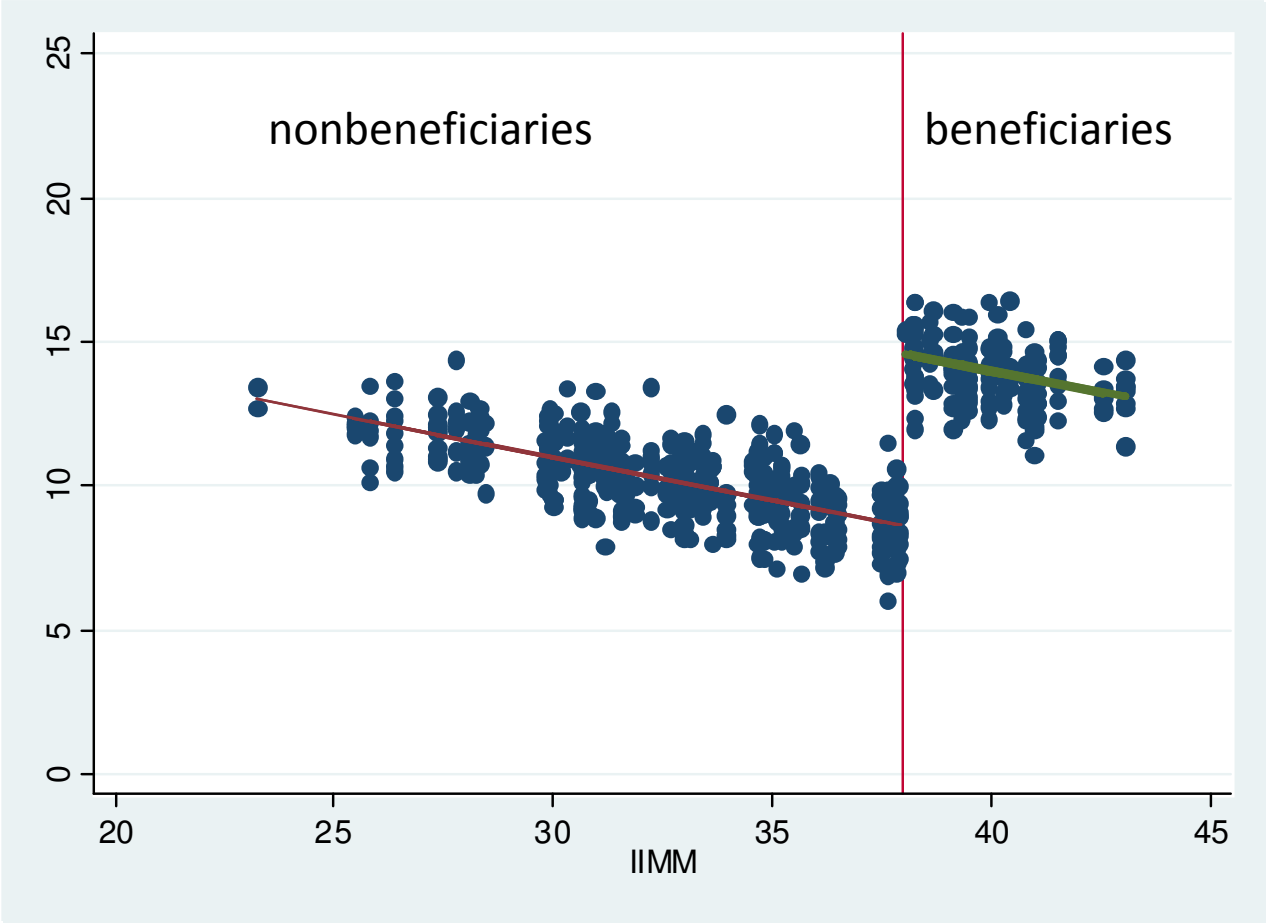
How RDD measures impact

- Before start of the program



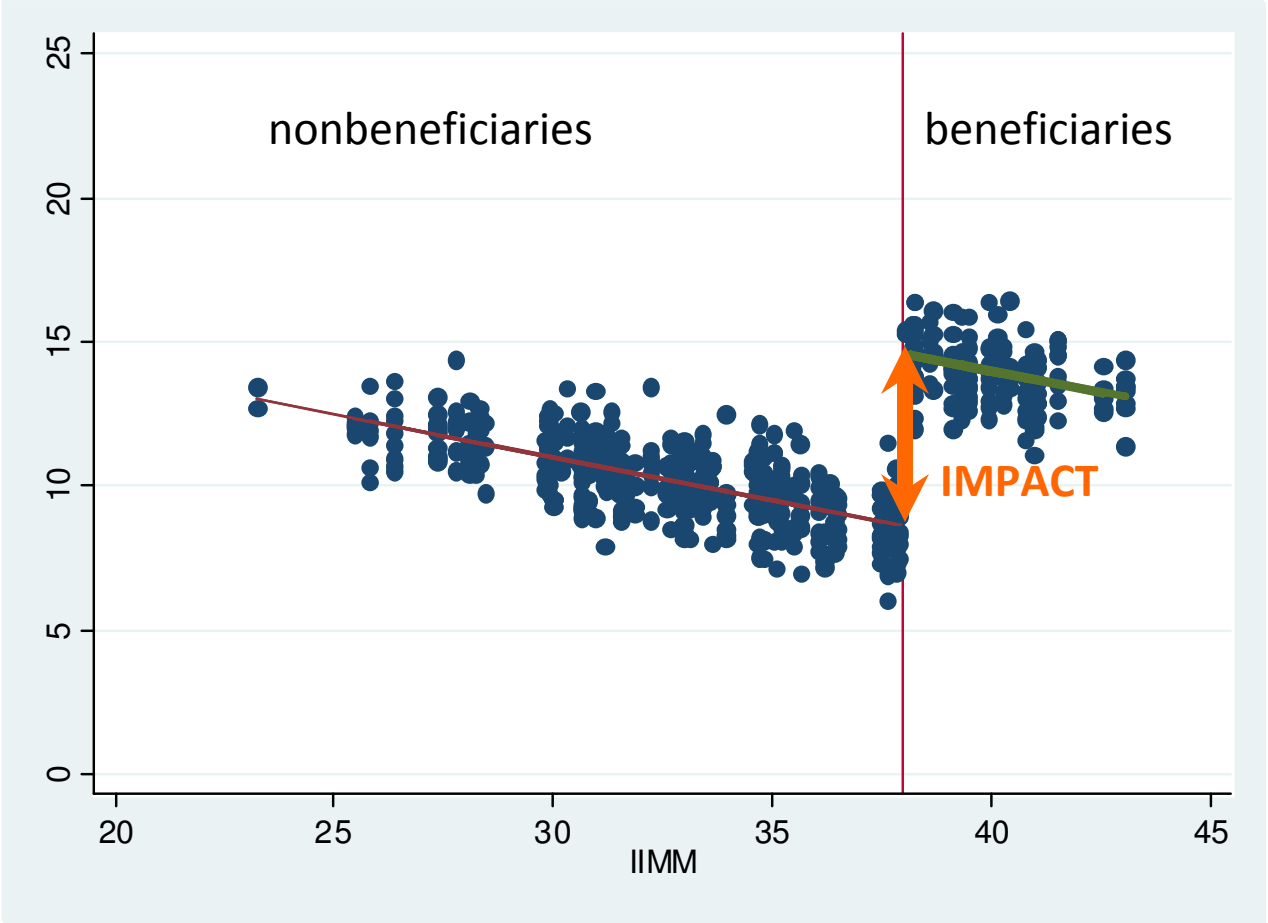
How RDD measures impact

- After the program



How RDD measures impact

- After the program



Regression Discontinuity Design

- RDD's strength lies in the fact that households lying either side of the discontinuity share common characteristics, yet those on one side are "treated" while those on the other are not

- The difficulty in applying RDD in the context of agricultural innovations is finding a suitable cut-off
  - de Janvry et al (2010) suggest using geographic boundaries that are not physically or politically important but this sensible restriction limits the opportunities to use RDD
  - Because RDD provides a LATE estimate of impact, it may be difficult to capture distributional effects (eg impact on the "poorest of the poor")

Matching

- Match beneficiary households to nonbeneficiary households with similar observable characteristics in the baseline survey
  - Matched nonbeneficiaries form a comparison group to represent counterfactual outcomes for beneficiaries
- Estimate impact as the difference in weighted average outcomes between beneficiaries and matched nonbeneficiaries over time
- With matching, the quality of the evaluation depends heavily on the quality of the matches: not as convincing as randomization
- Matching methods based on cross-sections make untenable assumptions about the distribution of unobservable characteristics across adopters and non-adopters
- Matching with panel data will account for fixed unobservables that affect “levels” but time varying unobservables remain a source of bias
  - Include retrospective questions at baseline to account for “trajectories” (eg changes in asset holdings that occur pre-intervention)
  - Include questions on time varying events (eg shocks) in follow-up
- Can be used to assess continuous treatments.

Six final thoughts

- ***There is an overwhelming case for basing SPIA-sponsored evaluations on double difference designs.***

- While randomized designs are powerful, they are not perfect. In the context of evaluations of agricultural innovations, special care needs to be taken to consider:
  - Ethical issues associated with withholding treatment
  - How actions by implementers and by subjects can confound interpretation
  - The trade-off between the "tightness" of the evaluation design and the extent of external validity

- Are we too pessimistic about measuring "unobservables"?
  - "I have the power to make important decisions that change the course of my life."

Six final thoughts

- Are we interested in the mean impact or the distribution of impact across certain household types? *(Pinstrup-Andersen)*

- The distributional effects (males, females, adults, children) within households are important.
  - I am very skeptical that increases in agricultural innovations that increase yields will significantly reduce undernutrition in child under the age of 24 months (b/c income elasticities are low and b/c innovations may have adverse effects on other inputs into nutritional status) but I am open to the possibility that agricultural innovations that reduce vulnerability to adverse environmental shocks may have positive impacts on undernutrition (*Hazell point about thinking about the economic model underlying the determinants of outcomes*)

- Impact evaluation must be complemented by process evaluation