Overview and evaluation of methods used in *ex-post* impact assessments of agricultural technology

Alain de Janvry, Andrew Dustan, and Elisabeth Sadoulet

Outline

1. Context and purpose of this report

• **Objective**: **Review advances in impact analysis methods that can be used to improve the *ex-post* impact assessment of CGIAR technology adoption**.

- Highlight fundamental **challenges** faced in epIAs

- Review **recent efforts** to address these challenges

• Focus is on **impact of adoption** of a technology **on outcomes**:

- Different from studies focusing on **impact of research expenditures** on outcomes.

- Different from analysis of **determinants / barrier to adoption**

Research → Technology → Adoption & Diffusion → E_S*E_C = Outcomes

E_S = Effect size = Estimated impact = ATT

E_C = Effect scale = Diffusion = Numbers of adopters or area

• For simplicity, focus on the more technical/biological CGIAR technologies, such as seed varieties, vaccines, etc. (not management practices or policy)

2. OVERVIEW OF EX-POST IMPACT ANALYSIS: OBJECTIVES AND CHALLENGES

Which technologies?

- **Yield-increasing and cost-saving** (e.g., some seed varieties)

  Challenges: impacts may be quite marginal compared to next-best;
  Seeds may be hard to tell apart (Gollin)

- **Risk-mitigating** (e.g., drought/flood-resistant varieties, livestock vaccines)

  Challenge: typically, adverse event must occur for benefits to be observed.
  Can take a long time to observe

- **Quality-improving** (e.g., nutrient-enriched varieties like QPM)

  Challenge: measuring impact difficult when market fails to assign a higher price to high-quality variety.

Short-run microeconomic versus long-run and aggregate effects

Main goal of epIA: Find total effect of a technology after some diffusion.

However diffusion fundamentally transforms the nature of the impact, so E_S is endogenous to E_C: **Cannot take E_S (ATT) as predetermined to E_C (diffusion)**.

Key points:

(i) **E_S is not static**: The average impact of technology for adopters changes over time and across farmers for many reasons:

- Learning effects: $\Delta E_S/\Delta E_C>0$
- Changing composition of adopters: more Schumpeterian farmers adopt first, $\Delta E_S/\Delta E_C<0$.

(ii) **The value of E_S changes with general equilibrium effects**: With diffusion, if aggregate supply increases and prices fall, there is transmission of more benefits to consumers and other sectors of the economy.

→ Sharp contrast in methods needed for micro vs. long turn aggregate epIA

• **Measure of microeconomic impact: <u>when can be done</u>?**

Main objective is to **measure the average impact for adopters (E_S)** based on comparison with counterfactuals.

But not all impacts of adoption can be measured:

Only feasible if general equilibrium effect not important:

- Relatively early adopters
- Substitution for an earlier variety without much aggregate supply effect; technology that applies to a small group or region
- Or even to the whole country if it is a tradable good (but labor, land NT)

Once diffusion is complete, there is no counterfactual left to measure impact

• **Long-term aggregate impact of a technology or string of technologies.**

After spillover and general equilibrium effects have taken place (at the limit no counterfactual anymore), **need very different methods / units of analysis**.

1. Econometric estimations: Need to observe the past, before technology adoption occurred

Foster and Rosenzweig : **Panel analysis** of small "economies" (villages) that have differentially benefited from technological change over time. Village and time fixed effects control for much of the potential confounding factors. Estimate the effect of yield changes on welfare, poverty, etc. Very demanding in terms of data, and econometric skills. Goes a long way, but still cannot identify the aggregate effect that applies to whole India (absorbed in the village and time fixed effects)

2. Simulations: **Using microeconomic measures**/estimations of impact (yield or TFP), and assuming a model for the functioning of the economy (sectoral supply and demand in the economic surplus methodology, many markets in CGE), and necessary elasticities. **Simulate** what would have happened (if *ex-post*) or might happen (if *ex-ante*) to welfare without these microeconomic changes.

Useful tool for re-scaling the microeconomic measures.

Report focuses on the short run microeconomic impact analysis

Why this focus?

- There have been many **recent advances** in methodology for impact analyses.
- The CGIAR centers **perform a lot of these**, that could be improved.
- Even when we want to know about long-run GE impacts, the short-run microeconomic impacts are often **used in the simulation models**. We however need to be really careful about the population of adopters over which E_S is measured vs. used for simulations.

Outcomes of interest: Choice of indicators

Profitability of a technology drives its adoption. It is thus natural to consider the impact of adoption on **farm-level restricted profits**, rather than **yields** which ignore adjustment of inputs.

Impacts on measures of welfare are also important:

Income

Expenditure

Poverty

These give an idea of how much adoption actually affected wellbeing.

But caution: increase in farm profits due to adoption of a technological innovation are **often not large enough** to substantially improve welfare (income, exp.) or lift households out of poverty.

3. MICROECONOMIC IMPACT ANALYSIS: ANALYTICAL FRAMEWORK AND KEY ISSUES

(1) Adoption decision

Farmers with access to a technology make the decision to adopt (takeup T) based on expected profitability of adoption.

Adoption by farmer i at time t (T_{it}) takes place when expected profits rise due to adoption:

$$T_{it}(Z_{it}, U_{it}, \varepsilon_{it}) = \begin{cases} 1 \; if \; E\pi^*(Z_{it}, U_{it}; T_{it} = 1) - E\pi^*(Z_{it}, U_{it}; T_{it} = 0) + \varepsilon_{it} > 0 \\ 0 \; otherwise \end{cases}$$

Z are observable characteristics, U are unobservables, and ε everything unrelated to the maximization problem that affects adoption. Adoption is assumed binary here.

(2) Impact of adoption

Outcome (Y) depends on observables (X), unobservables (V), adoption (T), and other random factors (η)

$$Y_{it} = Y_{it}[X_{it}, V_{it}, T_{it}(Z_{it}, U_{it}, \varepsilon_{it}), \eta_{it}]$$

This is just a **Heckman** selection model

Identification beyond reliance on functional form requires exclusion restrictions (IV): some Z that predict T are not part of the X that explain Y (but, very difficult to find without design)

Estimating the effect of adoption on adopters: ATT, not ATE

The model predicts that the group of adopters will be different from the group of non-adopters. In particular, those with the highest expected profitability from adoption will adopt most frequently → **Selection into adoption**

Thus the **average effect of adoption for a random household** (the average treatment effect or **ATE**) is probably very different from the average effect of adoption on those who found it favorable to adopt (the average treatment effect on the treated, or **ATT**).

Impact analyses need to make sure that it is the **ATT** being estimated because this is the effect that is actually realized. The ATT need not be larger than the ATE: households with highest potential returns from adoption might not have access to it due to various constraints (e.g., Suri for Kenya).

Selection and the counterfactual

Households **select into adoption** on the basis of both observable and unobservable characteristics. If we have two **observationally identical** farmers and one adopts while the other does not, there is a good chance that they differ importantly on **unobservable** characteristics.

If the unobservables that affect adoption are correlated with the unobservables that affect the actual outcomes, then estimating the two equations above will give **biased estimates** of the technology's impact on outcomes.

Examples of unobservables that affect both adoption and outcomes:

- Farmer's ability, entrepreneurship
- Soil quality
- Weather shocks that occur early enough to affect technology adoption

"Selection on observables" is unlikely to hold when selection into adoption has occurred—need research designs that allow for unobservables to differ between adopters and non-adopters.

Addressing spillovers from adoption

Adoption by one farmer might affect the outcomes for other farmers, both adopters and non-adopters, even without GE changes in the economy:

- Local employment and **wage** effects
- Local effects on input and output **prices** if limited tradability
- **Learning**-from-other effects—own adoption raises other farmers' returns to adoption in same social network
- Environmental **externalities**

Average impact of the technology is not just the ATT. It is the (**ATT + average spillovers**) from adoption: spillovers are important contributors to impact.

4. Current approaches to microeconomic impact analysis: deconstruction of use of farm trials, PSM, and Dif-in-Dif

(Dec.1) Problems with experiment station and on-farm trials (to measure E_S)

- Do not allow inputs and management practices to **adjust endogenously**

  Neither imposing no change in complementary inputs, nor imposing the optimal package correspond to what adopters would do

- Do not get **ATT for actual adopters**

  Randomized treatment gives ATE, because participants to the experiment include producers / plots that **would not have adopted** in the real world. Not necessarily a lower bound on ATT: could be that households with highest potential return are somehow more constrained in adopting.

- Experiment station trials cannot directly **estimate outcomes** of interest (e.g., farm profits, incl. labor costs, etc.) **beyond yields**

(Dec.2) Problems with selection on observables designs

- **Ubiquitous** in the literature: not so much OLS, but Propensity Score Matching (PSM)

- Rely fundamentally on assumption that no **unobservables** determining adoption are correlated with the **unobservables** that affect outcomes (omitted variable bias).

- **PSM does not alter this assumption**, it just makes the selection equation more flexible w.r.t. observables.

- Basically, **PSM does not work when choice has been exercised by non-adopters**: at equality of observables, they evidently differ in what made them opt out of adopting (same in microfinance)

- Typical approach: Cross sectional survey, observable determinants of adoption, PSM for counterfactual, "impact of adoption" → Does not work

- Rosenzweig at UCB: "**Time to put to rest** using PSM in adoption-impact studies!"

(Dec.3) Difference-in-differences (with observed or matched adopters/non-adopters)

- **Welcome** addition to portfolio of methods

- But few have actually used it in agricultural technology epIAs due to data requirements: **need panel data** for adopters and non-adopters.

- Relies on assumption of **common pre-adoption trends** among adopters and non-adopters– testable with long panel but rarely have information for that.

- **Lesson learned**: need to plan the evaluation in advance to get baseline data. But if we're planning in advance, maybe we can do better and have an explicit research design?

(Dec.4) Addressing spillovers

- Most studies compare adopters to non-adopters in the same village

- If adoption affects non-adopters, then the **counterfactual group is contaminated** and no inference can be made

- But if we compare adopters in one village to non-adopters in another village without access to the technology, we need to account for **clustering of standard errors at the village level**. Two-village (or few-village) comparisons won't work.

- Unit of analysis is the village: **need a sufficient number of villages** to meet power of test requirements, which is demanding

5. Conclusions: From deconstruction to reconstruction

1. ATT (impact of adoption, E_S) is **endogenous** to adoption/diffusion, not predetermined: varies over time due to learning, entrepreneurship, GE effects

2. To measure outcome as E_S*E_C, must be **very careful** that E_S corresponds to E_C:

- E_S from **experimental** plots/farm trials does not work: exogenous T gives ATE instead of ATT; outcome measure (yield) insufficient
- E_S from **PSM** does not work: matched C different from T on key non-observables that explain adoption (≡MFI)
- E_S from **DD** can work, but requires panel data to verify prior parallel trends

3. Impact can be measured for **limited adoption** with local spillovers, not GE effects and not if diffusion complete

4. Unit of analysis is village if there are **spillovers**: demanding for power of test

5. Should use **prior design** instead of ex-post recuperation: RCT or natural experiment (S-side rollout, discontinuity design, other ideas specific to case!)

Next!