Annex 3: Measuring the poverty impacts of agricultural research

Workshop organized by SPIA, University of Minnesota, 26 July 2014

Overview and main messages from the final session

Coming the day before the start of the annual American Applied Economics Association (AAEA) meetings, this one-day workshop attracted over 60 participants, mostly by invitation from SPIA. The meeting was attended by SIAC Program Steering Committee (PSC) members Greg Traxler and Julian Alston. The agenda was designed in such a way that we would hear briefly (10 mins per presentation) from a small group of presenters (case study leaders of SPIA-commissioned studies + other relevant papers) that have used a particular approach to assessing poverty impacts. In each session this was followed by a group of discussants that were briefed with the task of critiquing the role of the broad family of methods for the purpose of assessing poverty impacts from new technologies, rather than dwelling too much on the details of the individual papers.

This format worked well but required some steering from the Chair, overseen by Doug Gollin. The take-home messages from the first four sessions (on observational methods, micro-models, RCTs, and macro-models) was that it was a very hard task to find strong evidence with a single study that will cover the long causal chain from new technology to long-run poverty impacts. The take-home message from the final session 5 was, that despite the challenges, producing this evidence should be a priority for the CGIAR. Karen Brooks, Jim Oehmke and Greg Traxler all highlighted the importance of poverty outcomes for their institutions, and Peter Hazell and Bob Herdt offered some optimism that these significant challenges may be overcome. Nonetheless, SPIA remains concerned about a mismatch between the consensus of the group of academics gathered in the first four sessions, and the expectations outlined in session 5. Where should we draw the line on rigour for impact assessments in the context of balancing the standards for top-level academic research with the pressing concerns for evidence from governments and donors to help understand the effectiveness of their investments in agricultural research.

With this tension in mind, SPIA is currently working on a paper that will summarise and critique the literature on agricultural research and poverty, as well as outline its perspectives for good practice in this area. This is being written by Doug Gollin with research assistance from Lilli Probst and a draft will be delivered by 8th September, and presented and discussed at the forthcoming ISPC meeting in Copenhagen, on 15th September.

Introductory Session

Peter Gardiner, welcomed participants to the workshop and gave some context to the need for it. How can the CGIAR articulate what we are doing in a convincing manner? Doug Gollin, explained that this workshop is primarily about methods, and that SPIA is open-minded and ecumenical across the spectrum of possible identification strategies. Each method has its own boundaries and limitations, and SPIA believes that it is important to uphold standards of rigour and to highlight what is and is not possible. Year by year monitoring of poverty changes from agricultural research has been suggested and SPIA is concerned about these expectations.

Session 1 – Observational micro studies

Melinda Smale, Jeff Alwang and Sosina Bezu presented their recent SPIA-commissioned studies that all employ variants on observational methods – instrumental variables, fixed effects methods, correlation random effects and control function approaches. Marc Bellemare cautioned that some of the “fancy” estimators used may actually give identification from functional form – that the identification may not really be there. In addition, testing for exogeneity of your instruments is not reliably achieved using the Hausmann test, for example. Where panel data exists, why not simply use difference in differences? The assumption of parallel trends is no less credible than some of the other identifying assumptions that the other estimators require. And in cases where there are more than two time period panel data to draw on, the case for difference in difference is stronger yet.

Michael Carter noted that poverty lines are arbitrary and not easily defended – in many cases the impacts for the poorest people may be negligible whereas those that are just below the poverty may benefit just enough to put them over the poverty line. In general, greater care is needed in handling heterogeneous effects along the full income distribution. The issue of continuous vs binary treatment was also discussed, as was the possibility of taking things out of the error term through methods such as generalisable propensity score matching. Can attributes such as "eagerness", or "entrepenurial-ship", which might otherwise be in the error term, be made observable, given its likely importance to the process of adoption? Andrew Dillon also emphasized that the search for a clear source of exogenous variation, a targeting rule that can be pulled out of the data, is preferable to allow for a regression discontinuity design to be implemented. He supported the idea that certain confounding factors, that would otherwise languish in the error term and potentially bias the estimations, can be mopped up by measuring "unobservables" such as "eagerness". Finally, it is important to understand whether the process being studied (e.g. the "choice" of variety) is an active choice from a range of possible options for the farmer, or is taking place somewhat by default (such as when farmers only have access to a single variety).–

In disscussion, Jeff Alwang wondered what the right amount of money is that the CGIAR should be spending on surveys and how this might be coordinated. Doug Gollin noted that SPIA's rough approximation is that 100 – 200,000 households are visited every year by CGIAR researchers, and so the benefits from greater coordination are potentially quite considerable. Tavneet Suri supported the calls on the part of the discussants to move back to linear models. Non-linear models are all biased in agricultural applications as errors are not independent and identically distributed (i.i.d). The control function approach, for example, will identify from functional form in the absence of an exogenous instrument. Michael Carter and others made a call for a greater focus on theory around the process of adoption and disadoption.

Session 2 – Micro-models

Stein Holden, David Raitzer and John Antle gave short presentations on their respective models, all different in formulation and applied in a diverse range of contexts from maize in Africa to rice in SE Asia to aquaculture in Bangladesh. Tim Dalton opened discussion by noting that mathematical programming represents a broad range of tools that are important across the research spectrum. In testing a hypothesis using models that are not inherently statistical in nature, there is a good deal of judgment required. There are different ways of achieving a degree of peer review over this process such as the example of the community of researchers using the Global Trade Analysis Project (GTAP) model, which for over 20 years have been sharing data and model extensions with each other. There can be a trade-off between flexibility of the model and the ability to characterise the context adequately. It is desirable if the model can be constructed in such a way that behaviour emerges rather than being imposed by the model, and the solution space has to be plausible.

Julian Alston asked the question "What should we try and measure most accurately?". The yield gap in agriculture is getting bigger, there are market failures in research, and importantly, agricultural research may not be most appropriate instrument for achieving the targets that have been set for it. Cheryl Doss emphasized that the models used to assess the poverty impacts of agricultural research should focus closely on the theory of change articulated for the research programs in question: "Why do we think this research will impact on poverty? How?" Models depend on data and they are only as good as the data we put into them. There are a number of measurement issues – relating to technology, adoption and poverty outcomes. Perhaps most importantly, there are serious concerns about heterogeneity of impact – that the elasticity varies for different kinds of household. For some questions this may not matter too much in the aggregate, but for poverty it is very important. Most fundamentally, if the agricultural research succeeded, what would the world look like? Land ownership patterns are likely to be part of this story – bigger farms and out-migration?

In discussion, Stein Holden noted that the model assumptions are becoming more flexible, away from strict rationality to include features such as procrastination and prospect theory. Doug Gollin asked the question of validation of different models using back-casting exercises – which of these models and for what contexts? John Antle noted that the Agricultural Model Inter-comparison Project (AgMIP) is trying to do these kinds of cross-validation and, importantly, simulates outcomes taking a more complete system's perspective and not looking at a single farm-level technology in isolation. Both Jerry Nelson and David Raitzer noted that poverty impacts could mean two things: poverty as the dependent variable, or focusing on poor people as the population of interest, zeroing in on what's happening to them. Both are important but can be different from another.

Session 3 – Randomized Control Trials (RCTs)

Michael Carter, Tavneet Suri and Kyle Emerick all presented work in progress using RCTs to assess the impacts of agricultural research. In general, the outcomes being studied in these experiments are short-run in nature and theory suggests that they correlate well with poverty. Mywish Maredia delineated the kinds of questions that can and cannot addressed using RCTs. For example, indirect effects over long time horizons are not well suited to this family of approaches. In the sense that RCTs are excellent at examining the on-farm constraints and complements to technology adoption, they can properly be viewed as research tools rather than tools for evaluation. Mary Arends-Kuenning rehearsed the arguments from the de Janvry review on RCTs that SPIA commissioned in 2011 – that experiments should be randomised at the community level and that the technologies should be offered at market price, not subsidised. The process of deciding on which RCTs to do and what doesn't get studied seems to be somewhat arbitrary. There is growing interest in studying the mechanisms through which treatment effects are studied – the Emerick paper is particularly convincing on mechanisms. Meena Meenakshi stated that she felt it was a good thing that the agricultural community have embraced RCTs – there has been explosion of them across multiple interventions and mediating factors. More effort on studying packages of interventions, picking up externalities, and examining cost-effectiveness would add to their value to agricultural research.

In discussion, Tavneet Suri expressed her opinion that the incentives are changing for researchers and donors – that more scale-ups are being funded based on the findings from RCTs. To learn effectively from RCTs, they need to be large scale and designed so as to pick up heterogeneous effects across different regions and target groups. In most cases, RCTs may be able to measure productivity changes in the short run, but unlikely to detect significant changes in poverty levels (Carter). On externalities, the RCT structure gives you the flexibility to go back and look at externalities if there are concerns. However, Kyle Emerick argued that looking beyond the producer is important at the stage at which the RCT is designed, in order to pick up wage effects etc. Alessandra Garbero thought that greater transparency around pre-analysis plans for effects that emerge later would be step in the direction of greater rigour. Similarly, greater emphasis on heterogeneous treatment effects would be appropriate. There was some agreement that the most useful role for RCTs may be in rigorously estimating parameters for feeding into macro models, particularly related to productivity effects.

Session 4 – Macro models

Xinshen Diao, Phil Pardey and Will Martin all gave short insights into their research in this broad field that cut across country-level models (Diao), spatially-explicit modelling (Pardey) and global macro models (Martin). Terry Roe noted that the share of workers in agriculture decreases with growth over time, and as incomes grow, the share of agriculture decreases. There are strong indirect effects via the rest of the economy. We need studies that tell us the story behind this process – what are the causes of the macro processes? George Norton emphasized the links between the micro and the macro perspectives. Ideally, one needs a team to analyse these things properly. Jerry Nelson asked about the CGIAR reform process and the lack of ex-ante analysis that informs these major changes of change currently underway. The comparative advantage of the CGIAR may shift and donors say that this should happen. But important questions – should the CGIAR do systems research? What is the role of gender research in

the CGIAR? How does biophysical research interact with economics? – are currently poorly supported by spatially-explicit modelling.

Doug Gollin steered the conversation in the direction of specific methods by asking when does spatial analysis matter? When are there important interactions between farmer decisions with spatial and temporal variations? There was wide agreement that it was necessary to work together to get better data and share it effectively. The results may not be consistent across micro, meso, and macro models – there can be different magnitudes or signs, and these should explored in collaborative groups. In that context, Phil Pardey pointed out the changing nature of the market for "innovation" – agricultural R&D representing about 5% of the global market of approximately $1.2 trillion USD annually.

Session 5 - "Reducing Rural Poverty" as a System-Level Outcome for the CGIAR system

This session focused on a major question: should the CGIAR retain rural poverty reduction as an SLO? Karen Brooks started by making the case for keeping poverty reduction as a target for the CGIAR. We should step back from it if we felt that either a) the CGIAR research does not impact on poverty or b) we can't measure it. However, she feels that neither of these apply and it would send a strong negative signal to donors if we scrapped it. The well understood and often-claimed impact pathway story (ag research → new innovations adopted → productivity improvements → poverty reduction) has changed and that story needs to be re-stated and revalidated, since it is playing our differentially now. Thus,the stylized facts of direct effects, food prices and employment and wage effects need an update. Globalisation, with more open economies and the cost of trade and communication decreasing has meant that many are sceptical of a poverty-reducing role for agriculture. However, we need better geographic differentiation in updating the Strategy and Results Framework (SRF) of the CGIAR.

Peter Hazell agreed in retaining poverty reduction as an SLO. He reminded the group that much of the discussion and debate about evidence of poverty changes from agricultureal R&D already played out in the 70s and 80s. Had some of the 'early lessons' from the GR been acted upon (e.g., large farmers benefited the most), some of the biggest impacts on the poor from the GR would have been missed, because small holders eventually did adopt and many indirect effects (price reductions of major staples) accrued disproportionately to the urban poor. Hazell also made the case for country-level CGIAR impacts, such as the Diao study, but in a manner that better integrates the micro- and the macro- perspectives. There are big challenges in using poverty (#s of poor) as a dependent variable in our micro models, especially when looking at the effects of adoption of a single technological innovation, given the typically small share of income contributed by a single crop and, typically, the small percentage change in yield/profits through adoption of a new variety. Thus, studying a single intervention may not show any significiant impacts, whereas measuring the direct and indirect effect of a composite of CGIAR related interventions taken together, in a given country, may well show significant results. Also, these things take time – the dynamics of change are important and the time lags can be significant. However, it is important to distinguish attribution of impact by the CGIAR from contribution by the CGIAR – the NARS and others play important roles. This would ideally be done with cooperation across CRPs and centers. The danger is that we get very rigorous using reduced form approaches but miss information about the pathways inherent to the more structural models used in the past.

Greg Traxler outlined how the Bill and Melinda Gates Foundation broadly subscribes to the "poor but efficient" model of a smallholder. Raj Shah in 2006 decided to direct a significant portion of the $31 billion endowment from Warren Buffett towards a new agricultural development unit. However, agriculture is competing within the Foundation against global health programs (50%) and US programs (25%). The agriculture research community has to convince Bill and Melinda regarding the distribution of benefits, differential effects, gender and poverty impacts and the number of people lifted out of poverty is an important metric in this argument. Thus there

continues to be very strong demand for poverty outcome indicators – it's required to keep the funding going (from BMGF); headcount of the poor is important, along with other outcome indicators that rely on a strong theory of change. Adoption of technologies is central to the argument for funding research – we need better numbers on this as well. As a community of researchers we need to ensure that RCTs are used to address the most important and relevant questions rather than being supply-driven. On the other hand, many of the macro models are difficult to explain and describe to non-specialists. Perhaps the piece of the puzzle that is missing most glaringly is the kind of institutional analysis that Vernon Ruttan made famous.

Jim Oehmke read out an official statement on behalf of USAID's Feed The Future (FTF) initiative which has the goal of ending poverty and hunger by 2025. FTF believe that it can and has had significant impact on reducing poverty, and has reported to the US Congress on the reducing numbers of poor people in countries where USAID works. The program will be renewed for a 2nd 5-year phase. FTF needs to know the numbers of poor people but is not wedded to a firm attribution to specific interventions within the overall portfolio of sustainable intensification and agricultural transformation. We need to shift from a focus on comparative statics to comparative dynamics to enable analysis of climate smart agriculture, resilience and related outcomes. All of these perspectives are within the context of CAADP and the mutual accountability that exists for the commitments therein. Oehmke noted that outcome measures such as gross economic surplus, or consumer/producer surplus are no longer of such interest to donors; they want to know about poverty reduction.

Bob Herdt sketched out a model for how the multiple pathways from technology adoption to poverty effects intersect and interact in specific contexts. Economic transformation is a process of accelerated change with agriculture becoming less and less important, within which significant changes take place in the demand for food. Perhaps it might be instructive to also take a different approach to those outlined in the previous sessions and start with poor people and work back to look at the role that technologies play in their lives? He argued for a better understaning of the poor and their livelihoods to put into perspective how agriculture can contribute to poverty reduction.

In discussion, Julian Alston pushed back on the idea of national accounting for CGIAR impacts – that it would represent a fallacy of composition. How to represent what is happening in the rest of the world at the same time? Doug Gollin noted that this is an example of a wider issue in international development policy – there might be other things that reduce poverty more than investments in agricultural research. Tom Tomich reflected that the concept of poverty used in the workshop was narrower than he might have expected – that asset control and capabilities had received scant attention, similarly the role of agricultural research as a human enterprise. Michael Carter endorsed this perspective of poor people as agents rather than failures and that this has emerged as an issue in the second generation of cash transfer programs.

Ruben Echeverria noted that we struggle to make good ex-ante predictions, but do we need the word "rural" in the wording of the System-Level Outcome (SLO)? Tavneet Suri argued that you can't reliably compare across instruments for reducing poverty – that there aren't datasets available that would allow you to do such a thing. Karen Brooks stated that she is happy to be identified as an advocate for agricultural research – that we have endured 17 years of divestment from agriculture which contributed to the food crisis. Peter Hazell returned to the theme of cross-CRP and center tracking of progress in perhaps 8 – 10 countries with the most significant CGIAR investments. Greg Traxler encouraged SPIA to keep pushing to measure poverty impacts, or the space will be "claimed" by the M&E folks and the work will be less rigorous. Jim Oehmke urged the CGIAR to make the case for those aspects of its portfolio that will make the most difference show how it fits in a broader development agenda. Bob Herdt's take-home message was that we should take seriously the portfolios of the poor and try and understand the lives of poor people .

Anita Regmi and Peter Gardiner offered closing remarks on the importance of the workshop in the context of the CGIAR reform process, on behalf of the Consortium and ISPC, and Doug Gollin thanked everyone for their active participation, highlighting that SPIA will draft a paper that reflects on the workshop and circulate it for comments in late September.

Annex 1 - Agenda

08.30 Welcome and Introductory Remarks (Doug Gollin, Peter Gardiner)

08.40 - 10.10 Session 1 (Paul Glewwe, chair) What can we learn from observational micro studies?

- Lessons from SPIA-commissioned studies (Melinda Smale, Jeff Alwang, Sosina Bezu; 10 mins each)
- Methodological challenges, identification problems, data needs and survey design (Marc Bellemare, Michael Carter, Andrew Dillon; 10 mins each)
- Open discussion (30 mins)

10.25 - 11.55 Session 2 (C. Ford Runge, chair) What can we learn from model-based approaches drawing on micro data?

- Lessons from SPIA-commissoned studies (Stein Holden, David Raitzer, John Antle; 10 mins each)
- Methodological and data challenges; model validation; concerns with heterogeneity (Tim Dalton, Julian Alston, Cheryl Doss; 10 mins each)
- Open discussion (30 mins)

11.55 - 13.05 Session 3 (Alessandra Garbero, chair) What can we learn from RCTs?

- Lessons from SPIA experience and discussions (Michael Carter, Tavneet Suri, Kyle Emerick; 10 mins each)
- Cost effectiveness; applicability to large-scale and long-term impacts; challenges (Mywish Maredia, Mary Arends-Kuenning, J. V. Meenakshi; 5 mins each)
- Open discussion (25 mins)

13.45 - 15.00 Session 4 (Shenggen Fan, chair) What can we learn from meso, macro and cross-country studies?

- Lessons from SPIA-commissioned studies and related research on the CGIAR (Xinshen Diao, Phil Pardey, Will Martin; 10 mins each)
- Issues of data quality and model quality, sensitivity to assumptions, identification, validation (Terry Roe, Jerry Nelson, George Norton; 5 mins each)
- Open discussion (30 mins)

15.15 - 16.45 Session 5 (Doug Gollin, chair) "Reducing Rural Poverty" as a System-Level Outcome for the CGIAR system

- What have we learned about documenting and measuring impact? Is impact assessment feasible? What can we hope to find? (Panel discussion: Karen Brooks, Peter Hazell, Greg Traxler, Jim Oehmke, Bob Herdt; 10 mins each)
- Open discussion (30 mins)
- Responses from panelists (2 mins each)

16.45 - 17.00 Closing remarks and wrap-up (Doug Gollin, Anita Regmi, Peter Gardiner)

17.00 - 18.00 Reception organised jointly with the University of Minnesota, Department of Applied Economics and Agricultural Education

Annex 2 – Participants

1	Tahirou Abdoulaye	IITA
2	Julian Alston	UC Davis
3	Jeff Alwang	Virginia Tech
4	Robert Andrade	Univ. of Minnesota
5	John Antle	Oregon State University
6	Mary Arends-Kuenning	University of Illinois
7	Aminou Arouna	Africa Rice
8	Aden A. Aw-Hassan	ICARDA
9	Katherine Baylis	University of Illinois
10	Marc Bellemare	Univ. of Minnesota
11	Sosina Bezu	Norwegian Univ. of Life Sciences
12	Karen Brooks	IFPRI
13	Marie-Charlotte Buisson	IWMI
14	Michael Carter	UC Davis
15	Tim Dalton	Kansas State University
16	Xinshen Diao	IFPRI
17	Andrew Dillon	Michigan State University
18	Cheryl Doss	Yale University
19	Ruben Echeverría	CIAT
20	Kyle Emerick	Tufts University
21	Shenggen Fan	IFPRI
22	Monica Fisher	CIMMYT
23	Steve Franzel	ICRAF
24	Alessandra Garbero	IFAD
25	Peter Gardiner	CGIAR ISPC Secretariat
26	Paul Glewwe	Univ. of Minnesota
27	Doug Gollin	University of Oxford
28	Guy Hareau	CIP
29	Peter Hazell	IFPRI
30	Bob Herdt	Independent Consultant
31	Stein Holden	Norwegian Univ. of Life Sciences
32	Meenakshi J. V	University of Delhi
33	Tim Kelley	CGIAR ISPC Secretariat
34	Lakshmi Krishnan	CGIAR ISPC Secretariat
35	Catherine La Rochelle	Virginia Tech
36	Ricardo Labarta	CIAT
37	Lucy Lapar	ILRI
38	Ravinder Malik	IWMI
39	Mywish Maredia	Michigan State University
40	Will Martin	World Bank
41	Sam Mohanty	IRRI (IAFP)
42	Jupiter Ndjeunga	ICRISAT
43	Jerry Nelson	Independent Consultant
44	George Norton	Virginia Tech
45	James Oehmke	USAID
46	Phil Pardey	Univ. of Minnesota
47	Frank Place	ICRAF
48	Lilli Probst	University of Oxford
49	David Raitzer (Skype)	Asian Development Bank
50	Anita Regmi	CGIAR Consortium
51	Byron Reyes	Michigan State University
52	Terry Roe	Univ. of Minnesota
53	Ford Runge	Univ. of Minnesota
54	Melinda Smale	Michigan State University
55	James Stevenson	CGIAR ISPC Secretariat
56	Tavneet Suri	MIT
57	Daniel Suryadarma	CIFOR
58	Tom Tomich	UC Davis
59	Greg Traxler	BMGF
60	Roberto Valdivia	Oregon State University
61	Ira Vater	CGIAR ISPC Secretariat
62	Di Zeng	Virginia Tech