Photo: Georgina Smith / CIAT

SIAC Program (2013 – 2017) Final Report

September 2017

CGIAR | Independent Science and Partnership Council

SIAC Program Final Report

SIAC, Strengthening Impact Assessment in the CGIAR, was a four and a half year program of work led by SPIA that aimed to broaden the coverage and widen the range of impact measures of CGIAR research, and help develop the capacity of impact assessment scientists within the System. The program started in 2013 and was motivated by a strong and growing demand across the development community for high-quality *ex post* impact assessment to support evidence-based decision making in the CGIAR. SIAC activities also aimed at developing new methodological tools for collecting the data the CGIAR needs for impact assessment, as well as building the capacity of the CGIAR centers and CRPs to do the same and, in the process, seek opportunities for institutionalizing the process of adoption and impact-related data.

The key objectives of the SIAC program were as follows:

- Objective 1 (Methods): Develop, pilot and verify innovate methods for collection and assembly of diffusion data;
- Objective 2 (Outcomes): Institutionalize the collection of diffusion data needed to conduct critical CGIAR impact evaluations;
- Objective 3 (Impacts): Assess the full range of impacts from CGIAR research;
- Objective 4 (Building a community of practice): Support the development of communities of practice for *ex post* impact assessment within the CGIAR and between the CGIAR and the development community more broadly.

There are a number of unique features in this work program. Under Objective 1 and 2, collaborations with parallel data collections efforts, such as by the World Bank Living Standards Measurement Survey (LSMS-ISA) and national statistical agencies were developed. This allowed SIAC to leverage high quality data collection efforts to test the appropriateness and feasibility of methodologies such as DNA fingerprinting to collect adoption data at scale. Another type of collaboration, developed under a competitive funding window, was between CGIAR researchers (social sciences and natural sciences) and external impact assessment experts to document the adoption of NRM innovations claimed to be successful, and understand the extent to which innovations were (or were not) taken up. Defining and implementing a cost-effective strategy to track adoption of CGIAR innovations at scale is an urgent task, which will require partnerships between geneticists, geo-spatial scientists, NRM researchers, and social scientists. Under Objective 3, a highly competitive grant-making process funded studies of economic, social, nutritional, and environmental impacts (see the Annex for a list of SIAC funded studies). All else being equal, having studies being led by external and independent researchers ensures a high degree of objectivity and credibility. Nevertheless, CGIAR researchers play a critical role, owing to their unique combination of technical skills grounded in their knowledge of specific geographies and/or institutions, and CGIAR researchers were involved to a varying extent in many of the studies.

The program, which had a budget of approximately US$ 12 million[1] over the four and a half year duration, ended on 30th June 2017. This report summarizes the work carried out, the results obtained and the implications of the findings. It is organized around the Objectives spelled out in the SIAC program of work.

Please note that the SIAC funded studies are currently out for external review, so results or interpretations may change over the coming months. All studies are or soon will be available on the ISPC website. Drafts are available on request to Lakshmi Krishnan (lakshmi.krishnan@fao.org).

1 Funding for SIAC for the period 2013 – 2017 comes from three donors:

USD 5,238,788 from the Bill and Melinda Gates Foundation (BMGF); USD 4,453,122 from CGIAR Window 1; USD 2,430,133 from the CGIAR ISPC

OBJECTIVE 1: Develop, pilot and verify innovative methods for collection and assembly of diffusion data (METHODS)

Underpinning this objective was the development of a robust set of methods for routinely tracking adoption of CGIAR-related technologies in a cost-effective manner. Such information is a prerequisite for achieving the highest quality assessment of outcomes and impacts. A set of activities are designed to test innovative ways of assessing the adoption of improved varieties of crops, livestock and fish technologies, and agronomic and natural resource management interventions, with the goal of eventually embedding protocols derived on these tests into large-scale surveys carried out by other institutions outside the CGIAR, such as the World Bank's Living Standards Measurement Survey – Integrated Surveys of Agriculture (LSMS-ISA). There are four main areas of activity here, much of which is managed by Michigan State University.

Activity 1.1. Advance methodologies for tracking the uptake and adoption of improved varieties

The objective of this Activity is to pilot test and validate alternate approaches to collect variety-specific adoption data against a reliable benchmark to determine which method/approach is the most cost-effective (i.e., which method provides a given level of accuracy at the least cost). The idea is to come up with 'lessons learned' and recommendations on methods / approaches that can be used in scaling up the collection and assembly of diffusion data on improved varieties. The following crop-by-country combinations were targeted:

1. ***Cassava in Ghana*** *(IITA / MSU)*
2. ***Maize in Uganda*** *(SPIA / LSMS-ISA / Diversity Arrays)*
3. ***Beans in Zambia*** *(CIAT / MSU)*

Two further crop-by-country combinations were added to the SIAC portfolio by SPIA at the start of 2015, as part of the collaboration with LSMS-ISA:

4. ***Cassava in Malawi*** *(SPIA / LSMS-ISA / Diversity Arrays)*
5. ***Sweet potato in Ethiopia*** *(SPIA / Diversity Arrays)*

Each of these studies followed a common format – comparing data collected using DNA fingerprinting (representing a benchmark) against a series of alternative survey-based alternative data collection methods (representing potential low-cost proxies for the benchmark that could subsequently be incorporated in other surveys). The results from these studies are summarized here – please note that a number of these papers are still under peer-review so check back on the ISPC website over the coming months for updates, and note that there could still be changes to the summary results pending the outcome of the reviews.

The table below shows comparisons of DNA fingerprinting results with aggregate adoption estimates from alternative data collection methods (MAPE is Mean Absolutely Percentage Error). Note that the correspondence for adoption rates of specific varieties is lower. Also, note that an aggregate adoption rate that is close to the benchmark may mask potentially large numbers of false positives and false negatives (i.e. farmers thinking they are growing an improved variety when it is actually local, and vice versa) in the data. The % share of farmers who are correct in the aggregate (for either improved or local), and at the level of specific varieties, is another way of conceptualizing the same results.

Table 1 – Summary results from five DNA fingerprinting studies in activity 1.1. For each study, a reference adoption estimate is established through DNA fingerprinting of plant material from a sample of households. At the same time as the samples of plant material are taken, a series of questions are asked to the farmer (columns labelled B – D). Column A shows the equivalent aggregate adoption estimate established through convening an expert opinion elicitation process (see Walker and Alwang, 2015 "Crop Improvement, Adoption and Impact of Improved Varieties in Food Crops in Sub-Saharan Africa" for details).

			Aggregate adoption estimate (i.e. % share of improved varieties)			
			Expert opinion	Farmer self-reported data from survey		
Study	Sample size	**DNA Reference**	A: Expert opinion on extent of adoption of improved varieties	B: "Is this an improved or local variety"?	C: "What is the name of this variety"	D: Phenotypic protocol questions
Cassava, Ghana (MSU / IITA)	495 farmers; 914 plant samples	**4%** strict	36% (DIIVA, 2010)	**6%** (50% MAPE)	**2%** (50% - 94% MAPE)	**0%** (100% MAPE)
Maize, Uganda – Grain (SPIA / World Bank / Diversity Arrays)	550 farmers; 550 grain samples	**100%** [e]	**69%** (31% MAPE)	**45%** (55% MAPE)	**44%** (54% MAPE)	**12%** (88% MAPE)
Beans, Zambia (MSU / CIAT)	402 farmers; 855 seed samples	**16%**	9.5% (DIIVA, 2010)	**13%** (19% MAPE)	**4%** (75% MAPE)	
Cassava, Malawi (SPIA / World Bank / Diversity Arrays)	1,200 farmers; 1,200 leaf samples	**13.5%**	61% (DIIVA, 2009)	**21%** (56% MAPE)	**0%** (100% MAPE)	**5%** (63% MAPE)
Sweet potato, Ethiopia (SPIA / World Bank / Diversity Arrays)	231 plots; 231 leaf samples	**63%**		**69%** (10% MAPE)	**89%** (41% MAPE)	**74%** (17% MAPE)

The results from the maize in Uganda study were sufficiently surprising that we opted to replicate the research in a second round of data collection in 2016. This replication was further complemented with samples obtained by mystery shopping for maize seed in all the agro-dealers in the Iganga and Mayuge districts. These agro-dealer samples are intended to help us understand whether the informality of the seed system is a significant mediating factor that increases the probability that farmers have incorrect information about the specific variety they are growing. DNA fingerprinting analysis of the second round of data collection is still ongoing, as we have augmented and replicated multiple samples in the reference library. Early results suggest that we have problems of a lack of genetic purity at both the highest level in the seed production system (replicate samples in the reference library of breeders' seed of the same variety are not sufficiently identical), as well as at intermediate levels.

It should be noted that the field and laboratory protocol for this study has been the subject of critique from CIMMYT scientists from two perspectives: our use of a grain crop-cut for DNA fingerprinting, rather than seed or leaf; and that the DArT methodology cannot account for the presence of three-way hybrids. We have rebuffed the first critique on the grounds that fertilization from outside the plot is likely to be minimal, and that there are many practical advantages to using grain from crop-cuts from the perspective of future scaling-up of these analyses. The second critique is currently being investigated before we proceed to publication.

Implications of these results:

All data collected in surveys are inevitably subject to measurement error, and it is the responsibility of social scientists to take all practical means possible to ensure that these errors are minimized. This set of DNA fingerprinting studies bring to light the extent of measurement error associated with our cumulated body of data on varietal adoption. We have a number of Asian case-studies that apply DNA fingerprinting as part of long-term large-scale impact studies (SIAC activity 3.1, studies on: lentils in Bangladesh; C88 potato in China) and show a closer correspondence between farmer self-reported data and the DNA fingerprinting results (see Table 2 below). It is a small number of studies to drawn strong inferences from, but we have evidence consistent with the hypothesis that varietal adoption data from surveys carried out in Asia are more reliable than is the case in Africa.

Table 2 – Summary results from 2 DNA fingerprinting studies in activity 3.1

			Aggregate adoption estimate (i.e. % share of improved varieties)	
Study	Sample size	**DNA Reference**	A: Expert opinion on extent of adoption of improved varieties	B: Self-reported data from survey
Lentils, Western Bangladesh (Virginia Tech/ICARDA)	1000 farmers; 1697 seed samples[a]	**45.4%**	69.11% (Panel constituted)	**49%** "Is this an improved or local variety"?
Potatoes, Yunnan-China (Virginia Tech/CIP)	141 farmers; 141 samples (88 leaves, 53 tubers)	**97%**	NA	**100%** "Is this C-88"?

Activity 1.2. Develop protocols for tracking diffusion of natural resource management technologies

MSU issued a call for pilot projects in July 2013 and two studies were subsequently commissioned:

1. ***Innovative use of mobile phone based applications in tracking adoption of Natural Resource Management Technologies in India*** *(CIMMYT)*
2. ***Hyperspectral signature analysis: a proof of concept for tracking adoption of crop management practices in Gazipur, Bangladesh*** *(IRRI)*

Final reports from both pilots were received in mid-2015 and the results were ultimately disappointing from the perspective of discovery of new data collection tools that could be ready to be scaled up. The CIMMYT study featured survey responses from a highly unrepresentative set of respondents, showing the biases inherent in phone surveys, as well as the constraints researchers would face in scaling up such an approach from the perspective of the assembly of a sample frame. The IRRI study was ultimately unable to establish a unique hyper-spectral signature for crop residues, and through having this work externally reviewed, we have understood more about the heterogeneity of remote sensing approaches and the strengths and limitations of different methods.

A further pilot study was added in 2015, outside of the competitive process (activity 2.4 on institutionalizing data collection):

3. ***Measuring adoption of conservation agriculture: A study in Ethiopia*** *(SPIA, World Bank LSMS-ISA)*

Frederic Kosmowski (SPIA Research Associate based in Addis Ababa) designed a study to test the following methods for collecting data on soil cover from crop residues – one of the pillars of conservation agriculture, and a source of significant controversy in Sub-Saharan Africa owing to poor data quality.

A. Enumerator asks the farmer to estimate the % soil cover in their plot
B. Enumerator uses photo aids to help the farmer estimate the % soil cover in their plot
C. Enumerator uses a drone to take an aerial photo of the plot which is then digitized and a % cover estimated
D. High-resolution satellite imagery
E. Enumerator lays a transect rope with knots which is used to estimate for the field (reference)

Fieldwork in five enumeration areas in Ethiopia took place between November 2015 and February 2016, with data collection covering 200 households (350 plots). Satellite imagery was obtained from a commercial provider following a tendering process. This study helps us understand the extent to which survey-based data on soil cover is accurate relative to an objective reference, and also whether using drones or satellite imagery are viable alternatives. Results suggest that survey-based methods tend to underestimate crop residue cover, compared to direct observation in the field. Where quantitative estimates of the extent of cover are needed, the accuracy of such survey-based estimates are improved by using a visual aid protocol. The work is the subject of a World Bank Working Paper, and has been published in Environmental Management. The findings led to the visual aid protocol being taken up by the Ethiopian Statistics Agency for forthcoming agricultural surveys.

Implications of these results:

Methodological innovation is risky and may not result in a data collection process that can be scaled up. In the case of the phone-based application, there is selection bias associated with picking up the phone and responding to questions (see also the IFPRI-Geo Poll study in activity 2.2). In the case of remote sensing, hyperspectral signatures may not be sufficiently distinct or stable to be scaled up as a method for tracking adoption, but other remote sensing approaches may be more promising (see the UC Santa-Cruz-Nong Lam; ICRAF; and IFMR-led studies under activity 2.2). Where data reliability can be improved through modifications to existing survey questionnaires the marginal costs of doing so are very low while the gains are potentially substantial, particularly for panel surveys with rich socioeconomic data (see activities with the World Bank LSMS-ISA in activity 2.4).

Activity 1.3. New institutional approaches to collecting technology diffusion data

Most diffusion surveys in the past have depended on CGIAR research teams, either working on their own or working in collaboration with national programs and statistical services to collect the data. In many countries, there are private market research firms as well as private survey firms engaged in carrying out household surveys for academic purposes. A call for proposals was issued by MSU with a focus on doing a case study in India. The call was issued in February 2015, and either for-profit or non-profit entities with the relevant capacity were eligible to apply. Three projects were commissioned following a competitive procurement process as follows.

Table 3 – Summary of the diffusion surveys commissioned under activity 1.3

	Survey firm	States (Districts) covered	Technologies	Farming system
1	Synergy	Haryana (Karnal) Bihar (Vaishali)	Zero-till, Direct seeded rice, Laser-land levelling	Wheat-Rice
2	CASPL	Haryana (Karnal) Punjab (Ludhiana)	Zero-till, Direct seeded rice, Laser-land levelling	Wheat-Rice
3	Nathan	Andhra Pradesh (Anantapur, Kurnool)	15 soil conservation measures promoted by ICRISAT	Groundnut-based systems

Further to these three pilot studies, MSU carried out their own validation surveys in the same districts, working with the same survey company across all three studies. The purpose of these validation surveys was to compare: a) the costs of the local enumerator approach used by the three service providers to a more "conventional" centralized validation survey; b) the difference in the statistical estimates obtained through the same. The results for each of these studies, alongside the results of the three commission studies, are summarized in the following two tables.

Table 4 - Adoption rates established through the local enumerator approach (Local En.) and a conventional survey for specific technologies in the rice-wheat system in different states in India (studies 1 and 2 combined). ** indicates statistically significant difference at p <0.05

		Laser land levelling	Zero till	Direct seed rice
Ludhiana, Punjab	Local En.	51.7	5.6**	0.8
	Conventional	48.8	0.3	0.3
Karnal, Haryana	Local En. 1	67.3**	18.5**	0.8**
	Local En. 2	72.4**	22.4**	0.2**
	Conventional	61	3.1	4.2
Vaishali, Bihar	Local En.	1.5**	1.4	0
	Conventional	0	1.1	0

Table 5 - Adoption rates established through the local enumerator approach (Local En.) and a conventional survey for specific technologies associated with groundnut production promoted by ICRISAT in Andhra Pradesh (study 3). ** indicates statistically significant difference at p <0.05

		Soil Bunds	Field Bunds	Broad Bed and Furrow	Contour Bunds	Land Levelling	Polythene Mulching	Nala plugs	Sunken Pits	Farm ponds	Masonry dams	Well recharge pits	Penning
Anantapur, Andhra Pradesh	Local En.	32.6	38.3**	2.0**	15.6**	38.5**	0**	0.4	0.1	12.5**	0.2	0	20.1**
	Conventional	37.4	1.2	59.6	0	7.3	8.9	0.6	0	1.3	1.2	4.5	44.9
Kurnool, Andhra Pradesh	Local En.	4.3	1.6	2.5**	0	5.7	0.2	0.1	0	3.6	2.1	0.4	69.3**
	Conventional	14	0.1	30.4	0	6.3	0	0.2	0	2.6	0.5	0	35.8

Implications of these results:

Collecting the same data in the same sites using different approaches can significantly change the summary statistics inferred from the data. Without a clear benchmark for data quality, it is hard to know the direction of the bias associated with either a local enumerator approach or a conventional survey – one cannot be assumed to be better than the other.

Activity 1.4. Develop and disseminate best practices for collecting diffusion data

The idea with this activity was to take stock of the methods used, results and lessons learned from activities 1.1, 1.2 and 1.3, with the goal of generating guidance for the CGIAR system more broadly. The Policies, Institutions and Markets (PIM) CRP partnered with SPIA and MSU in organizing and participating in a workshop on 3rd and 4th August 2016 in Boston (immediately after the AAEA meetings).

With such interesting and complex findings from the DNA fingerprinting work, and the broadly disappointing pilots for other alternative data collection methods trialed, the process of writing a guidance document was deferred until such a time as more evidence is available to help shape that document. As can be seen below for activities 2.2 and 2.4 (as well as for a number of the impact evaluations reported in Objective 3), there is a lot of rich material on measuring adoption and diffusion that is coming available from elsewhere in SIAC - we have sought to weave in a methodological development approach into the way we implement the other activities. A series of methodological communication products are in the pipeline for late 2017 and throughout 2018.

OBJECTIVE 2: Institutionalize the collection of the diffusion data (OUTCOMES)

The objective here was to compile and make available the best information on outcomes that are at least plausibly attributable to CGIAR research outputs, and on a large-scale. This is where a key bench-marking function for the CRPs is most obviously fulfilled by this program. Large gaps in existing adoption databases for genetic improvement technologies (activity 2.1), natural resource management technologies (activity 2.2) and policy-oriented research (activity 2.3) were to be filled for priority regions. In addition, under activity 2.4, the World Bank Living Standards Measurement Study-Integrated Surveys of Agriculture (LSMS-ISA) team and SPIA and Centers worked together with NARS partners and statistical agencies to see how some of these processes could best be integrated into existing surveys to reduce cost and increase frequency of data collection. MSU explored similar objectives in Zambia, Mozambique and India.

Activity 2.1. Organize the collection of crop germplasm improvement research related direct outcomes

This activity has built on the DIIVA and TRIVSA projects, and focus on the collection of varietal diffusion data in South and South-East Asia. MSU led a process in which varietal release and varietal adoption data were collected for 134 crop x country combinations (CCCs) across 15 countries using expert opinion elicitation methods. The table below provides an overview of all the data now available following these activities. As for the DIIVA dataset, these data are being prepared to be shared on the ASTI website.

Table 6 – Summary of the varietal release and adoption data collected in South and South-East Asia in Activity 2.1

Crop	Crop Code	Number of varieties	% of total N	Number of countries (or states / provinces)	Total crop area covered under the targeted CCC	Total area across countries included (M ha)
Rice	1	5,445	35.98	22	83.7	125.4
Maize	2	4,945	32.68	25	45.8	57.8
Wheat	3	2,683	17.73	18	57.5	67.6
Potato	10	668	4.41	18	7.48	8.6
Sweet potato	11	603	3.98	16	3.59	3.9
Groundnut	5	380	2.51	7	3.98	5.8
Cassava	9	150	0.99	9	4.12	4.1
Barley	4	120	0.79	4	0.47	0.67
Lentil	8	73	0.48	2	1.32	2.13
Chickpea	6	55	0.36	2	1.34	1.3
Pigeonpea	7	11	0.07	1	0.60	0.61
TOTAL		**15,133**		**134 CCC (15 countries)**	**210**	**277**

MSU (in consultation with SPIA) had identified four CCCs for a validation of adoption estimates derived using expert elicitation method and/or secondary data sources. Two methods were used for validation estimating adoption using representative farmer surveys and DNA fingerprinting on all or a sub-set of seed samples.

The four CCCs identified for validation of Activity 2.1 are:

- Cassava in Vietnam
- Rice in Indonesia
- Wheat and lentil in Bihar

Table 7 – Comparison of expert opinion to DNA fingerprinting and survey-question summaries

Study	Sample size	DNA Reference	A: Expert opinion on extent of adoption of improved varieties	B: "Is this an improved or local variety"?	C: "What is the name of this variety"
Cassava, Vietnam (MSU / CIAT)	949 farmers; 1570 plant samples	**87%** strict **99%** inclusive	**95%** (4% - 8% MAPE)	**97%** (2% – 11% MAPE)	**10%** (89% - 90% MAPE)
Rice, Indonesia (MSU / IRRI)	809 farmers; 798 seed samples	**78%** strict **96%** inclusive	**92%** (5% - 18% MAPE)	N/A	**78%** (0% - 19% MAPE)
Wheat, Bihar (MSU / CIMMYT / ICRISAT)	3,278 farmers; 1604 seed samples	** pending*	**100%**	**65%**	**61%**
Lentil, Bihar (MSU / CIMMYT / ICRISAT)	1,000 farmers; 700 seed samples	** pending*	**43%**	**9%**	**9%**

Implications of these results:

What is striking from these results is the comparison with DNA fingerprinting results from the cases in Africa (activity 1.1). First, the aggregate level of adoption for these crops is estimated to be significantly higher than for the crops in Africa. Second, there is much closer correspondence between the DNA fingerprinting results and those of the other methods, notwithstanding the range for the "true" genetic result presented – an artefact of an incomplete reference library. Despite having 137 accessions in the reference library in the case of the Rice in Indonesia study, there are still 18% of the field samples that went unmatched to any variety. The fingerprinting results for wheat and lentil are not yet available.

Activity 2.2. Organize the collection of natural resource management (NRM) research outcomes

Following a delayed start after this work was transferred back to SPIA from MSU in 2014, SPIA finally issued a call for Expressions of Interest in October 2015, for case-studies focused on a number of priority NRM practices – country combinations we had identified through systematic searching through Center/CRP annual reports over a ten-year period (2005 – 2014). From this call, 62 expressions of interest were received, and these were reviewed internally by SPIA in November 2015. Proponents from 18 expressions of interest as well as a number of resource people and SPIA secretariat members were invited to participate in a workshop in Rome in December 2015 comprising: discussions of the nature of the priority practices; the existing data infrastructure in place in the relevant countries that can serve as the basis for generating adoption estimates; prospects for remote sensing; and group work clustered around the six practices. The overall objective of the workshop was to try and broker collaborations across interested parties to ensure we got a strong set of full proposals.

A further competitive process for funding followed the workshop and full proposals from an invited set of teams were externally reviewed (in a process led by consultant Paul Vlek) in January 2016. From this process, the following 9 studies were commissioned – the technology, country, method(s) of data collection, and teams are listed in turn, along with a brief description of the method and the main results.

1. ***Conservation agriculture in Mozambique and Zambia*** *(CIMMYT and ICRISAT)*

Mozambique and Zambia have been promoting conservation agriculture in a concerted manner since the 1990s. This study comprised a combination of workshops (to establish valid definitions of CA in each context) and summary analysis of existing data through meta-analysis. The definition of CA is rather diverse among the proponents in the countries, and development projects tend to be rather flexible, which may be optimal to the local context but that makes measurement and comparison across contexts difficult. Partial adoption (farmers doing only one of the component practices and/or only using CA practices on a sub-set of their suitable land) exceeds full adoption in the data reviewed in this study but the sum is still less than 10% of respondents. Where adoption is observed, it is primarily in areas where projects have promoted it. Furthermore, the meta-analysis carried out by this study team laid bare the problem of CA being a complex technology and one that is often incompletely or imprecisely defined in impact studies. Modelling drivers of adoption based on the meta-analysis findings supports the conclusions of the larger FAO-led study (below), with profitability, climate risk perception, and promotional activities being key in adoption by farmers.

2. ***Conservation agriculture in Malawi*** *(NMBU)*

Conservation agriculture in Malawi has been promoted through the lead farmer concept for a number of years. This study combines a sample of 180 lead farmers with a sample of 455 of their followers, as well a new round of long-run panel of 317 households. Adoption of the full CA package (minimum soil disturbance, permanent soil cover, rotation) in Malawi is negligible at less than 1%.

3. ***Conservation agriculture across East Africa*** *(FAO)*

This study uses a number of rich databases of surveys conducted twice (either as panel or in repeated cross-section) that included questions on CA technologies used by farmers in Zambia and Malawi. It shows that the full CA package, as defined by FAO, is rarely practiced by farmers (less than 1%). The components are more popular, particularly legume intercropping and rotations (3 – 50 %). The combination of one of those with minimum tillage is less than 5%. With the adoption of legumes and crop rotation, it is valid to ask whether this is actually simply traditional practice or if it represents the adoption of new technology. Geo-referenced CA and household data is linked with a series of bio-physical databases on climate and soil, and used to feed econometric models to examine the factors that covary with the adoption of CA components. The results show that a number of factors help predict adoption of these practices, among them are rainfall and rainfall variability and soil fertility. The key result from this study is that CA adoption as a package is negligible, and that the use of its components may not so much be adoption of technology but rather tradition, as a way of coping with the vagaries of the semi-arid tropics.

4. ***Conservation agriculture and fertilizer micro-dosing in Zimbabwe and Niger*** *(University of Illinois and ICRISAT)*

This report explores the adoption of micro-dosing as a fertilizer application technique in Niger. Using the nationally representative LSMS-ISA data, the study team calculate that only approximately 12 percent of farmers in Niger use any sort of chemical fertilizer. Since nationally representative data on fertilizer application technique is not available, the study relies on household survey data collected from randomly selected households in regions where crop production is most feasible. Here we see much higher fertilizer use rates of around 43 percent (see table 8 below). Of these fertilizer users, only about 18 percent apply fertilizer via micro-dosing. A large share, about 28 percent, use a spot application method that involves mixing fertilizer directly with the seed at planting – a method that is inconsistent with ICRISAT recommended practice.

Table 8 - Fertilizer application methods across regions in Niger (2014), including variants on micro-dosing (MD). (Note that the numbers in the columns do not sum to 100. This is due to overlap among some farmers how use more than one technique of fertilizer application in a given season.)

Variables (%)	Dosso	Maradi	Tillabéri	Zinder	Entire Sample
Any fertilizer	61.46	28.26	43.56	27.73	42.77
Any type of MD	18.38	19.88	13.99	20.36	17.95
MD pure	1.85	3.11	4.48	1.61	2.75
MD unburied at the side of budding plant	16.00	13.42	8.67	17.15	13.79
MD buried at the side of budding plant	0.69	5.51	2.42	3.1	2.65
Broadcast application	1.69	0.72	4.48	0.92	2.06
Line spread application	0.3	0.96	1.21	0.34	0.60
Fertilizer mixed with seeds	49.85	11.85	31.06	9.43	28.5

In Zimbabwe, the study team used survey data from 2012/13 to show that 40 percent of households in Zimbabwe use fertilizer while 34 percent of all households have used MD at some point. Among households that use fertilizer, over 80 percent apply fertilizer via MD for maize and sorghum, while over 60 percent apply fertilizer via MD for pearl millet. This suggests that the method of spot application of inorganic fertilizers has become a standard practice among smallholder farmers in the semi-arid regions of Zimbabwe, although work to establish how nationally representative these figures (for example, by benchmarking observable characteristics of the household against other surveys known to be nationally representative, such as the DHS) has yet to be completed.

Finally, for documenting CA adoption in Zimbabwe, a combination of three existing surveys were analyzed and found to be in agreement that farmers implement all components of CA on less than 10 percent of all farmed plots. However, even this might be an overestimate of adoption - the panel data included in the study (which ran annually from 2006/07 to 2010/11) shows evidence of a strong disadoption trend, finding that only about two percent of farmed plots can be considered CA plots based on the conventional definition of all three practices together.

5. Conservation agriculture in India, Institute for Financial Management and Research (IFMR)

This study focused on the Indo-Gangetic plain from the Punjab to Bihar. CA was de-constructed into three categories: partial zero till (ZT) practiced one season only; full ZT with < 30% ground cover; and full ZT with >30% ground cover. As proposed, 12 districts were selected with 240 villages in them in which 3600 farmers were interviewed. Even by this relaxed definition of full CA (missing the requirement rotation with a legume), adoption of full ZT and >30 % ground cover was insignificant. Using the more relaxed partial CA definition, adoption rates were below 3%, with the exception of Punjab where this was found on 16% of farms or land. In Bihar, farmers were not aware of the CA technology and did not adopt at all. In UP and Haryana, farmers were aware of the technology but adoption was very low. Only in the Punjab was it the case that farmers who were aware also seemed to have adopted partial ZT. This adoption rate was confirmed to a reasonable margin of error by remote sensing for the Punjab where ground truthing data were collected. Outside of the ground-truthing areas the remote sensing technology overestimated ZT adoption rates, suggesting that more extensive ground-truthing is needed per area of interest. The remote sensing technology will need considerable development before it is ready for application. The question remains whether it is interesting to pursue this remote sensing research to establish that the CA technology is slow to catch on.

Table 9: Estimate of Net Cultivated Area under partial zero-tillage

State	Net Cultivated Area (ha)	% of Net cultivated area under Partial Zero-Tillage	Area under Partial Zero Tillage (ha)
Punjab	3,965,749	15.96	632,933
Haryana	3,638,875	2.12	77,144
Uttar Pradesh	17,507,954	0.91	159,322
Bihar	6,134,254	0.00	0.00
Indo-Gangetic Plains	31,246,832	3.37	869,400

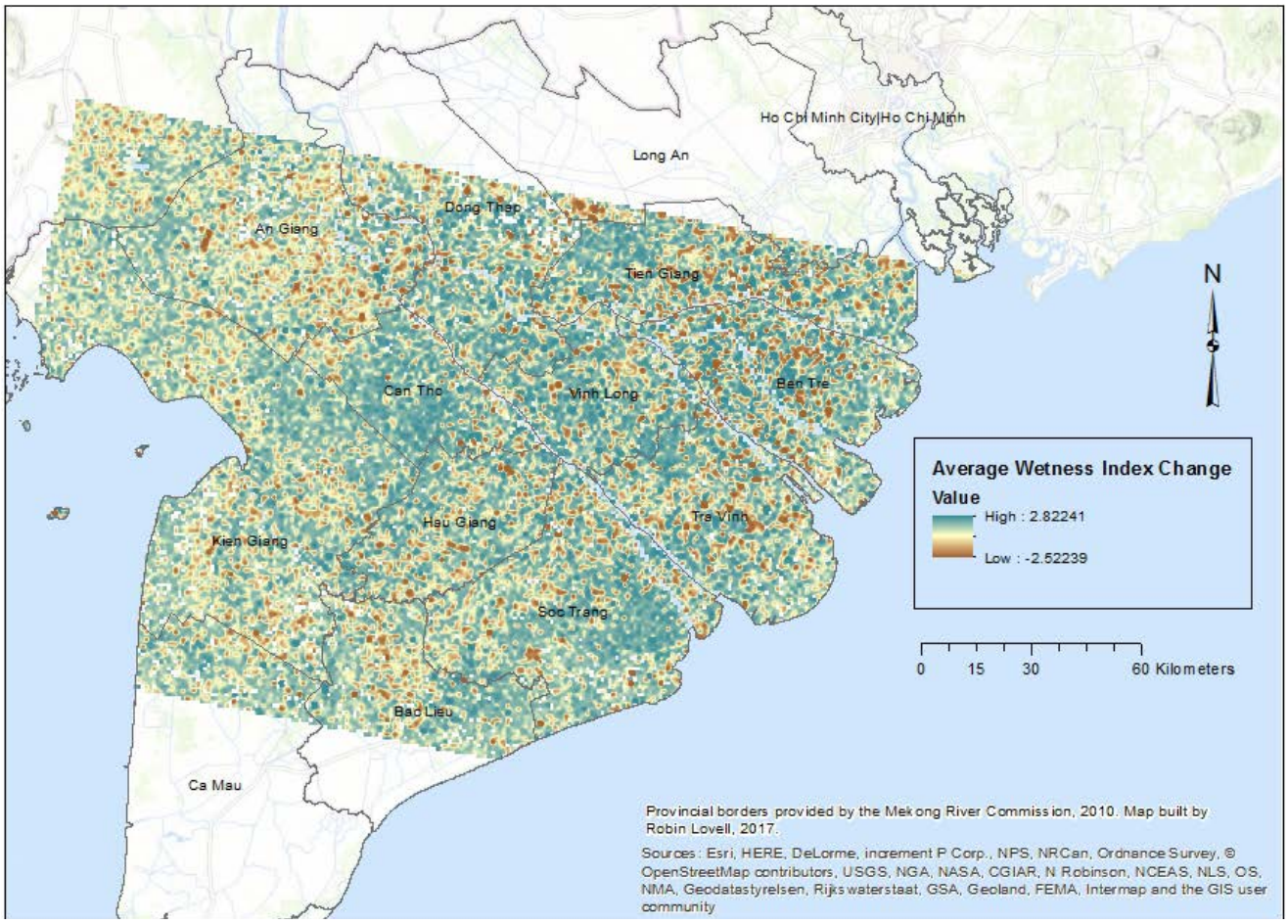
6. Conservation agriculture in Mexico (CIMMYT)

This study was closely linked to the previous study (#5) and aimed to test the potential for remote sensing to be used to estimate zero-tillage adoption. The specific remote sensing method employed was a technique developed at University of Leuven, using Sentinel 1 imagery. The European Space Agency agreed to adjust satellite settings for a three month to generate the appropriate imagery. The project did fieldwork in three states in Mexico targeted to get ground-truthing data. To date, the accuracy of the method has been tested on one state only so the project is still in progress. The study defines zero-tillage as farmers planting into the residue of the previous crop at least every other crop, and refers to this as CA. This definition of CA is convenient but confusing when comparing adoption information across projects. With this caveat, the results so far hold promise that remote sensing can be used to identify fields with zero-tillage, provided the Satellite imagery is available for the relevant dates when the impact of zero-tillage is visible. This work is planned to continue with national partners with the aim to document zero-tillage adoption or disadoption at a relatively large scale.

7. Alternate Wetting and Drying in Vietnam (University of California Santa Cruz and Nong Lam University)

This study focuses on irrigated (dry season) rice production in four provinces of the Mekong river delta in which AWD adoption was most likely. The team also looked at AWD during the wet season by incorporating the practice of field draining in rainfed flooded fields. IRRI had promoted the use of a perforated pipe for farmers to use in deciding on whether to irrigate, but this is not a necessary (easily observable) condition for AWD adoption as farmers quickly learned to use surface soil characteristics instead. The project characterized these conditions with a wetting index. A first outcome of the study is a map of AWD adoption in Vietnam (below), but further work on interpretation is pending. The remote sensing was complemented with a household survey in nine communes or wards, even if unfortunately the two exercises could not be linked.

Figure 1 – Map of "Average wetness index change" value, a proxy for the likelihood of adoption of AWD for Vietnam



8. Fertilizer trees and fodder shrubs in Zambia (ICRAF and Penn State)

The study uses geospatial methods and a farm survey to assess uptake of key agroforestry species, focusing in particular uptake of fertilizer trees across larger areas. In a change from the original proposal, the study focused on F. albida, (as other species of interest are considered not to be common in the region). It's important to point out however that F. albida is mostly growing naturally in farming systems in Farmer Managed Natural Regeneration (FMNR) and it is unclear if any organization can claim credit for this practice. The findings show that geospatial methods do indeed have high potential for mapping of key agroforestry species, and hence also for assessing uptake of fertilizer trees across larger areas. The ground-truthing was undertaken to verify the satellite model. A farm survey was done to assess the uptake of agro-forestry practices and a "process tracing" exercise was conducted to help in attributing the uptake to ICRAF/CGIAR. The survey showed 15% of cultivated fields with F. albida with 15 trees per ha. Agroforestry systems are virtually absent in Zambia today, which contrasts to the optimism of just 10 – 15 years ago regarding the potential spread of fertilizer trees across East and Southern Africa.

9. Integrated Soil Fertility Management (IFPRI, GeoPoll and IITA)

ISFM is defined in this study as a set of co-applied soil fertility management practices that include the use of improved germplasm, mineral fertilizers, and organic inputs combined with the knowledge on how to adapt these practices to local conditions. Earlier studies had shown that adoption of ISFM was low in Eastern Africa (<6%). This study primarily aimed to get a better estimate of this adoption using cell-phone interviews through partnering with a private sector company - GeoPoll. Ultimately, this attempt failed as the self-selection process into the sample of respondents (i.e. those who actually picked up the phone) was dramatically age-biased and much of the technological information was not credible. The project then examined panel data from Kenya and Zambia, but those surveys only included some of the necessary practices that defined ISFM. Finally, some recent national survey data from Kenya, Zambia and Rwanda were analyzed to assess ISFM adoption. The

traditional survey showed 9% adoption in season March-June 2012/13 and 5% adoption for season July-Sept 2012/13.

Implications of these results:

This set of studies has demonstrated, for the first time, some sobering low adoption rates for specific high-priority NRM technologies (conservation agriculture, alternate wetting and drying, agroforestry, integrated soil fertility management) for surveys carried out between 2012/13 (at the earliest in our set) and 2016/17 (for new primary data collected in the case-studies). The countries we have focused on with this call were chosen because they had been the subject of specific claims of widespread adoption in earlier CGIAR center annual reports. As such, the results provide an important reality check regarding the returns to NRM investments.

This evidence will need to be complemented with evidence on the effectiveness of NRM landscape level interventions. James Stevenson represented SPIA at a workshop in Cairns, Australia in June 2015, which concluded that too little attention is paid to demonstrating whether, and under what circumstances, a landscape scale approach is beneficial and will bring about impact. A paper in Sustainability Science summarized the consensus of the attendees.

Activity 2.3. Organize the collection of policy-oriented research outcomes

The goal of this activity is to compile and make available to CGIAR stakeholders the best available information on outcomes to which CGIAR research outputs have, at least plausibly, made a contribution. In particular, the activity builds an inventory of CGIAR policy-oriented research outcome claims that have been externally vetted by the Science Council as a part of the Performance Management System activities between 2006 and 2010.

Inventories were created for two periods, 2006-2010 and 2011-2014. For the first period, a consultant (Mitch Renkow) drew on earlier CGIAR PMS data files. Claims were reviewed (internally and externally) and ultimately 67 satisfied the specific criteria to be rated as "strong." For the second period, for which no PMS information was available, Renkow reviewed websites, annual reports and other materials to identify 27 claims. These were vetted by the consultant based on additional information provided by the Centers/CRPs. The database now comprises 104 research outcomes that bear on policies, including investment strategies and institutions, plausibly linked to Center / CRP outputs. However, very few of these claims have been assessed for impacts, and the final report identifies nine notable policy-oriented impact assessments conducted by the Centers, the research activity for almost all of which occurred prior to 2006.

Table 10 - Strong claims of policy influence, by period and type of claim

Type of policy influenced	2006-2010	2001-2014
Policy (trade, price, export ban, NRM, climate change, safety net, genetic resources)	4	19
International treaty	6	2
Regulation (Plant genetic resources, trade, patents, land use, markets)	15	3
Operations and maintenance (guidelines, methods)	28	2
Donor priorities	4	1
National investment priorities	8	-
Convening	2	-
Total	67	27

In addition, Renkow authored and presented a paper on 'Assessing the impact of policy-oriented research in the CGIAR: methodological challenges and reasonable expectations' at the International Conference on Impacts of Agricultural Research – Towards an Approach of Societal Values (French National Institute for Agricultural Research INRA, Paris, November 3-4, 2015). The paper offers a critical assessment of efforts by the CGIAR and kindred national agricultural research institutions to evaluate the welfare impacts of policy-oriented research conducted under their auspices.

Activity 2.4. Long-term institutionalization of collection of adoption data

SPIA's long-term vision in achieving this objective is to involve a broader and more diverse set of national institutional partners in the collection of adoption data so as to systematize the collection of nationally representative data (on a regular basis) in the most cost-effective way possible. MSU worked in India, Mozambique and Zambia to explore the integration of technology adoption data into existing surveys. On a parallel track, SPIA has been working with the World Bank Living Standards Measurement Study – Integrated Surveys of Agriculture (LSMS-ISA) team through two research associates – Frederic Kosmowski and John Ilukor.

***1. India** (MSU)*

From the visits and discussions with the heads of ICAR, NCAP, and the Department of Economics and Statistics for the State of Odisha, and the desk review of questionnaires used to collect different types of data through surveys that are routinely conducted in India (such as the crop cut experimental data, input surveys, agriculture census surveys and NSSO surveys), the emerging conclusion is that: India is a data rich country. There is an impressive amount of data being routinely collected (many at representative scale), and all these efforts are already institutionalized within the government system. However, despite these efforts, the fact remains that it is not easy to get an overall representative picture and trend of the adoption of different types of agricultural technologies that are generated by the Indian research system (and the collaborating CGIAR centers). Three potential constraints or reasons why this is the case are:

- Not all the data that is collected is even entered: From visiting Odisha, Mywish Maredia discovered that, due to lack of staff resources, and the short turn-around time they have to enter the data after they are collected, the data entry system is designed to only take numeric data and does not allow any string variables. Due to this practical constraint, the 'name' of the variety as reported by the farmer is never entered. Furthermore, government law on confidentiality of data prevents the data collecting agencies to share the actual survey or electronic data with anyone.
- Only aggregate level / processed data are available for public consumption: For the same data confidentiality reason mentioned above, the Unit level data for many of these surveys (i.e., input survey, crop-cut experiments, and ag census surveys) are never made public or available for research purpose. Only the aggregate level, processed data are made public in the form of Reports, which generally have a lag of 3 or more years between when the data are collected and when the reports are made public.
- Some surveys don't collect detailed data on specific technologies or data collected is not at a level (i.e., plots) that is needed to estimate the level (i.e., area) of adoption of specific technologies. For example, some surveys collect information on the adoption of IPM practices, use of inputs, type of seed etc. But these are mostly binary questions at the household level, which only allows estimating the number of farmers using certain practices or inputs, the 'extent' or 'intensity' of adoption.

Implications of these findings:

Due to these characteristics of the way data are collected, processed and reported in India, there is limited utility of these data for tracking technology adoption at a representative scale. There is certainly room for improvements in this data system, but a local institution or a research center needs to champion this cause. Someone with institutional connection with the key leadership positions at the agencies that have the mandate to collect the agricultural data needs to play an important role in bringing these issues to the attention of right people, and engaging them in discussions that can lead to structural changes in the type of data collected, the process of data entry, and making the data available for analysis to at least the researchers within the ICAR system. The goal would be to make some changes in the institutionalized data collection system so that the

data collected using public resources can serve the research and monitoring needs of the agricultural research communities.

***2. Mozambique** (MSU)*

The National Statistics Institute (INE) has delegated the powers of producing official agricultural statistics to the Directorate of Economics and Statistics (DEST) within the Ministry of Agriculture and Food Security (MINAG). One of the surveys routinely (annually, bi-annually or once every 3 years) conducted is the Integrated Agricultural Survey (IAI) which is representative at the provincial level. MSU has served as a technical partner to MINAG in conducting this survey (and its predecessor TIA survey) for decades. To ensure these surveys collect useful data that can help in tracking technology adoption by farmers, M. Maredia had requested a copy of the survey instrument to review and provided detailed suggestions for integrating some technology specific questions in different sections of the survey. These suggestions were shared with DEST through the local MSU staff based in Mozambique. We have been recently informed that in 2017 the plan is to do a 'light' version of the INE survey (i.e., a survey with fewer modules), and thus, DEST is not able to integrate all the suggested modifications in the questionnaire. However, some suggestions were taken on board and will be implemented in the planned survey this year. There was also an expression of interest from DEST to test new methods of tracking adoption of varietal technology, especially using DNA fingerprinting. However, no concrete plans were discussed and no activity was funded by SIAC to support this activity.

3. Zambia (MSU):

The Ministry of Agriculture & Livestock and Central Statistical Office in Zambia conduct Crop Forecast Surveys (CFS) annually that are representative of the small and medium scale holdings at the country level. This survey involves collecting data at two points in time --soon after planting (around March-April) and after harvest (September-October). M. Maredia had contacted the coordinator of this survey early in the year and requested a copy of the survey instrument to review for its content and ability to track technology adoption at the nationally representative scale. Based on this review, suggestions for modifications and addition of a one page section on the adoption of conservation agriculture technology were made to the CFS coordinator. The questionnaire was pilot tested in February 2015, but due to time constraint (the survey was already taking two hours), it was not possible to add any more questions in the first round. However, the survey team has agreed to integrate the one page suggested section focused on conservation agriculture technology adoption in the second follow-up round (postharvest survey). This survey is planned to be implemented in September for the small and medium scale holdings in Zambia. The proposed one page module was reported to be implemented in Zambia as planned. However, when a copy of the survey questionnaire was requested, it was not apparent that the suggested modules were in fact included in the data collection efforts.

Implications of these findings:

These studies show the difficulty of attempts to affect change in operations and culture in an organization (i.e. a national statistical agency) from the outside. We have learned the importance of long-term local partnership with the relevant government partners including the presence of team members in the country, so that trust is built and they can come to understand our need for these data. These lessons have fed directly into our activities with the World Bank LSMS-ISA in Ethiopia, Uganda and Malawi.

***4. Ethiopia** (SPIA and World Bank LSMS-ISA)*

The third wave (2015/16) of the Ethiopia Socioeconomic Survey (ESS) presented an opportunity for integrating a number of questions related to the adoption of CGIAR-related agricultural technologies. The ESS is a nationally representative survey of 4,000 households, and is managed by Central Statistics Agency (CSA) via a network of some 300 resident enumerators. SPIA were able to incorporate additional adoption-related questions into the ESS for the following technologies: Orange-fleshed sweet potato; Awassa variety sweet potato; Crop rotation in previous three years; Treadle pump; Motorised pump; Desi / Kabuli type of chickpea; Weather index insurance; Broad-bed maker; Improved livestock feed module. The data for the 2015/16 ESS survey wave was released in July 2017 and is available from the World Bank LSMS website.

5. Uganda *(SPIA and World Bank LSMS-ISA)*

The Annual Agricultural Survey (AAS) is a new survey funded by the Ugandan government and implemented by the Ugandan Bureau of Statistics (UBoS). The survey instruments were pre-tested in the second season of 2015. Plans for full implementation of the AAS survey have been delayed by UBoS, but SPIA were able to incorporate questions into the AAS for the following technologies: bean varieties; cassava varieties; maize varieties; sweet potato varieties; sorghum varieties; agroforestry; livestock; conservation agriculture.

6. Malawi *(SPIA and World Bank LSMS-ISA)*

In Malawi, the Integrated Household Survey 4 (LSMS-ISA panel survey) took place in the growing season 2016/17, with a number of questions incorporated by the National Statistical Organization following input from John Ilukor and James Stevenson.

Implications of these findings:

Greater specificity of agricultural technologies coded into large, well-institutionalized surveys, provides the potential for future adoption and impact studies to have a surer footing. However, there are still a number of data quality questions (e.g. in relation to varietal identification) that may require the introduction of specific additional survey rounds to be addressed as there is a limit to how much we can piggy-back off the core survey (for both length and achieving a representative sample reasons). This will be a major focus of any prospective phase 2 of the SIAC program.

OBJECTIVE 3: Assessing the full range of impacts from CGIAR research (IMPACTS)

While work under Objectives 1 and 2 paved the way for future ex post impact assessment studies, Objective 3 activities focused on carrying out a number of impact assessments of CGIAR research and development initiatives along the entire chain of causation - from research investments to the System-Level Outcomes. Since this causal chain is long and complex, SPIA approached it from a number of different perspectives: case studies that focus on measuring the impact of CGIAR research on health and nutrition (activity 3.0); long-term large-scale studies of impact for major areas of CGIAR investment (activity 3.1); sets of micro-scale impact studies using experimental methods (activity 3.2) to provide evidence on the impact of CGIAR research-derived technologies to adopting households; studies of a number of under-evaluated areas of research, e.g. irrigation and water management; livestock, agroforestry and biodiversity (activity 3.3); a system-level meta-analysis of ex post IA of CGIAR research (activity 3.4).

All studies commissioned under Objective 3 were solicited over two rounds (EoIs and full proposals) and were subject to independent external review. An inception workshop was held for each call, after the successful projects had been selected, at which proponents presented their plans and specific suggestions were made by SPIA, external experts and their peer study leaders, on every aspect of the study – from indicators to power calculation issues and econometric strategy.

Activity 3.0. Assessing the impacts of agricultural research on nutrition and health

This activity comprised a competitive two-stage call for proposals, led by Erwin Bulte at Wageningen University. The objective of the call was to generate evidence of the impacts of a range of agricultural technologies on nutrition and health. Five studies were commissioned in late 2013, and these are reported on below.

1. ***Adoption of high iron bean varieties in Rwanda*** *(CIAT, Harvest Plus, Virginia Tech, Rwanda Agric Board)*

This study aimed to combine an adoption study of high iron bean varieties and nutritional outcome measurement, based on full dietary intake data. The study specifically focused on areas where adoption was expected to be high – the power calculations for the research design assumed an adoption rate of 50%. In implementation the study team found lower levels of adoption (21 % of bean-producing HHs growing any HIB) and lower intensity of adoption than expected, with adopting HHs only planting an average of 7.7 kg (median 3 kg) of HIB in the 2015 season B. The nutrition module was dropped from the survey when it became clear that low adoption meant that the study would be unlikely to find any impacts on micronutrient intakes. DNA fingerprinting (which was planned for accurate varietal identification) could not be carried out owing to bureaucratic limitations imposed by the government, despite the collection of bean samples.

This is a disappointing outcome as we were hoping to have a study combining adoption and efficacy together. Yet the study entails important lessons as the “breaks” in the impact pathway are a concern for proponents of biofortified crops.

2. ***Shortening the hungry season through NERICA in Sierra Leone*** *(IPA, MIT, Sierra Leone Agr Res Inst)*

This study is a randomized control trial of the New Rice for Africa (NERICA), with different treatment arms representing the rice variety offered at different subsidy levels (0% i.e. market price; 50% and 100% subsidy for seed), crossed with a training intervention on how production practices should be adjusted to the new requirements of the new seed. The research team measured anthropometric indicators of nutritional status in children in treated and control households. Take-up for NERICA among the market price treatment group was 21%, rising to 62% take-up at 50% subsidy and 97% when offered NERICA free of charge.

The planting week was unchanged by the NERICA treatment whereas onset of harvest was earlier by five weeks for both the treated only and treated and trained groups. This is significant as the harvest being earlier comes at a time of acute food shortage in Sierra Leone – the hungry season. Yields for the group that was treated (received NERICA) and trained increased 23% relative to the control group, whereas the group that received NERICA seed but no training did not experience a yield benefit. This result is important as it demonstrates how

farmers can be deterred from persisting with potentially beneficial technologies that have high learning costs associated with them – without training, it may take a lot of trial and error on the part of the farmer to truly understand how the new technology could or should be used. Previous studies by the same authors have shown that NERICA is susceptible to crop failure when not grown under correct agronomic conditions, and these findings would suggest that farmer training may be a necessary condition for achieving certain development outcomes with NERICA. It should be noted that the yields observed in this study are very low at only 480 kg / ha in the control group. This is much lower than FAOSTAT data for Sierra Leone, but is consistent with the 2011 Agricultural Household Tracking Survey.

For the comparison between the outcomes for the experimental treatment group that received the NERICA seed free (100% subsidy) vs the control group, the study finds strong positive impacts on anthropometric measures – weight for height (wasting) and body mass index for age Z-scores – that persist up to the beginning of the next hungry season, though there is no statistically significant impact on middle upper-arm circumference. Impacts for the group with improved access but no training are positive, but not statistically significant and much smaller than those on the treated and trained group at the end of the hungry season, in line with the yield results. Importantly, the endline results show that the magnitude of the gains in anthropometric outcomes for households with improved access and training persisted increased over time, perhaps indicative of households expanding area under NERICA-6. Overall, the study provides strong experimental evidence of the impacts of NERICA rice on nutritional outcomes.

3. ***Crop diversification for food and nutrition security in Malawi and Ethiopia*** *(CIMMYT, Lilongwe University, Georg-August-University of Gottingen, Ethiopian Institute for Agricultural Research)*

This study applied econometric methods to data from Malawi and Ethiopia, to look at the effects of diversification of the cropping system on nutrition indicators. The methodological approach is quasi-experimental, applying a multinomial endogenous switching regression approach to cross-sectional data in the case of Malawi, and with a fixed effects panel structure applied (with data from 2010 and 2013) in the case of Ethiopia, to control for selection bias. In the Malawian case, the authors find that more diversified farming systems are associated with more diverse diets, but that the effects are small. In the Ethiopian case, the adoption of cropping system diversification (CSD) and improved maize varieties are associated with improvements in child stunting indicators, per capita consumption of calorie, protein, and iron; and a measure of the diversity of food consumed in the household. The study reported 15% less stunting for children in households adopting CSD and improved maize varieties, compared to 2% lower stunting rates for children in households adopting only CSD, and a 12% lower stunting rates for children in households adopting only improved maize varieties. The authors argue that these findings are consistent as hybrid maize growers are able to purchase highly diversified diets. These results support the intuition that, in rural areas in Sub-Saharan Africa, diversity in production may correlate with diversity in consumption but that this relationship is complex – certainly diversity in production is not a necessary condition for having a diverse diet. The relationship between diversity of production and consumption is often asserted but rarely tested, so these studies are a useful contribution. However, it should be noted that cropping system diversification is endogenous and it may be difficult to implement programs that provide exogenous stimuli that could increase crop diversification, thus limiting the potential policy implications of these results.

4. ***Looking beyond income: impact of dairy hubs on nutrition in Tanzania*** *(ILRI, Emory U., Tanzania NARS)*

This study applied difference in difference and propensity score matching together to examine the impact of dairy business hubs in Tanzania on a range of outcomes including women's diet quality. The authors report a negative relationship between women's dietary diversity and household participation in a dairy hub. While concerns related to endogenous dairy hub placement and low statistical power limit the interpretation of the quantitative results, they are nevertheless worrying and merit further analysis.

5. ***Nutritional impacts of irrigated horticulture in Senegal*** *(Columbia., George Washington U., MDG Center)*

This study aimed to piggy-back on a program to extend a combination of drip irrigation, nutrition messaging and specific packages of technical inputs to women in Senegal to grow vegetables in small-scale gardens. Baseline data were collected and summarized, and a trial of the nutrition messaging has been completed. However, the underlying program that was to be the subject of the evaluation has been stalled in bureaucracy

between the Senegalese government and donors. We still expect the program and associated evaluation will eventually take place, however we have no certainty on that.

Overall comments on the studies in activity 3.0:

Despite an attempt to link up CGIAR impact assessment focal points with nutrition scientists, this was a somewhat disappointing set of studies. This experience has laid bare the challenges of interdisciplinary work. In some cases, where the nutritional input was strong (e.g. Harvest Plus for the Rwanda high iron bean study; Johns Hopkins for the Senegal horticulture study) unrealistic expectations about adoption or program implementation prevented drawing lessons regarding nutritional impacts. It is perhaps not surprising that we see failure rates in implementation of impact assessments when there is a push for greater rigor – challenges in field implementation, methodological challenges of accurate ex-ante diagnostics and assumptions about adoption rates, etc. Longer time frames for studies could enable researchers to learn and adjust. At the same time, the NERICA study shows that important rigorous evidence can be generated on both impacts and mechanisms of CGIAR technologies with experimental designs that are powered and designed to test key hypothesis. These lessons are worthy of serious reflection in SPIA and beyond.

Activity 3.1. Long-term / large-scale impact assessment studies

Through this call, SPIA aimed to fund studies that sought to measure the adoption and impacts CGIAR research-related innovations that are believed to be widely adopted. Documenting adoption and estimating the direct and indirect impacts from widely-adopted CGIAR-related technologies and policies is of special relevance to CGIAR donors and stakeholders, particularly in a climate of high accountability and expectation of linkages between agricultural research investments and socially desirable outcomes.

Seven studies were funded out of the 12 full proposals received (8 impact + 4 adoption studies) in January 2015:

1. ***Adoption and diffusion of C88 potato variety in China: Spatial variability of productivity gains and cost savings and value chain development*** *(CIP, Virginia Tech, and Yunnan Normal Univ)*

C88 potato variety (resistant to late blight), developed under a cooperation between CIP and Chinese researchers, was released in 1996 and spread widely among farmers in Yunnan province. This study of 616 households in 47 villages showed that 23% of farmers were growing C88 in 2015 corresponding to 400,000 ha. Seed potato producers provided about 1000 tons of C88 out of the 50,000 tons total seed, and the survey found that most farmers use their own saved seed. Partly as a consequence of seed recycling and the resultant seed degeneration, decreasing yields and lower resistance to late blight, adoption is now lower than it has been in the past. Interestingly the seed system currently does not produce a high quantity of C88 seeds (only 1000 tons out of a total 50,000 tons of potato seeds) – the question is whether the seed system ever did produce a high quantity, or whether this adjustment occurred once demand (or lack of it) became evident. An expert panel estimated the time path of adoption between 1996 and 2014 and together with data from the study and other sources, and this was used to calculate a total of approximately $2 to 3 billion USD in economic benefits from C88 over the 19 year period. Furthermore, this study used DNA fingerprinting for a narrow application – samples of plant material of what farmers stated were C88 were collected for fingerprinting to confirm that they were indeed genotypes of that variety. The correspondence between these DNA fingerprinting results and the farmer statements were extremely high at 97%, suggesting a very low level of false positives. The sampling strategy does not allow to estimate the fraction of false negatives in the survey areas (farmers who are in fact growing C88 when they don't realize it).

2. ***Estimating improved Tilapia adoption using DNA fingerprinting: Philippines and Bangladesh*** *(WorldFish)*

In this adoption study, ninety-nine hatcheries from Bangladesh and 104 from the Philippines, representing about half the hatcheries in each country, were surveyed. Tissue samples from broodstock fish from the surveyed hatcheries were collected for DNA analysis using 1300 SNP markers which had been identified using the DArTSeq platform. The assignments made using the SNPs show some marked discrepancies with what hatchery operators believed they had – both in Bangladesh and in the Philippines. However, when grouped

into Genetically Improved Farmed Tilapia (GIFT), non-GIFT and GIFT-derived categories most individual fish were assigned to their expected category although 20-25% of fish were allocated to an alternative group. This was found in both Bangladesh and Philippines hatcheries. In aggregate, this study found that genetically improved tilapia from WorldFish (and fish strains derived from them) accounted for almost 53% of production from Bangladesh hatcheries and 40% from Philippine fish hatcheries. About half the farmed fish production in each of the countries in 2015 comes from such varieties.

3. ***Adoption of improved lentil varieties in Bangladesh: Comparison between expert estimates, nationally representative farm household survey and DNA fingerprinting*** *(ICARDA and Virginia Tech)*

Improved lentil varieties developed cooperatively by Bangladesh researchers and ICARDA have become widely grown as a "relay" crop in ripening rice fields, mostly in rainfed fields, since 1991. Lentil growing districts in the Western Bangladesh region comprise about 74% of national lentil area. The area planted to improved lentil varieties increased from 72,600 hectares in 2007 to 145,500 hectares in 2015. At the same time, a number of older varieties remain popular (for instance, BARI-3 released in 1996), and the study did not find evidence of improved yields in post-2006 varieties, in comparison to the varieties they replaced (which were also improved). Farmers were asked what lentil varieties they grew, when they started to cultivate each, why they chose the varieties, and why they didn't cultivate the others. Quality of newer varieties (post-2006) – not amenable to consumption preferences and poor post-harvest quality – are identified as some of the factors influencing varietal turnover. The study also compared varieties identified by the farmers during the household survey with DNA fingerprinting results: this showed that 89% of the samples were matched by the two methods. Using a combination of endogenous switching regression and instrumental variables approaches, the study team econometrically estimated that yields of improved lentils released since 1996 are more than 380 kg/ha higher than pre-1996 varieties (primarily traditional). Aggregated across the country the new varieties had an increased value of production of $26 million in that year.

4. ***A systematic and global assessment of the impact of CG technologies on poverty*** *(IFPRI and World Bank)*

This study applied a computable general equilibrium approach to assess the link between CGIAR research and poverty. A very challenging set of tasks were involved if this study were to be successful – the team had to assess the extent to which improvements in productivity have increased output and lowered costs in developing country agriculture; estimate the extent to which investments in CGIAR R&D have raised productivity in developing country agriculture; and then assess the implications of this productivity growth for poverty reduction in developing countries. The first two of these steps are extremely limited by data availability on a large scale, so the research team ran a "multi country, multi sectoral, dynamic computable general equilibrium model (MIRAGRODEP)" backwards through time to estimate productivity growth for each commodity based on data on prices and total output quantities. The step of parsing out the CGIAR R&D contribution to productivity growth from contributions from other sources relied on a Delphi process for eliciting estimates from experts. These various steps are inherently speculative, and SPIA had advised against pursuing this line of enquiry – rather, we would have preferred a systematic exploration of the parameter space that drives different relationships. Nonetheless, taking the results at face value suggests that productivity gains generated by agricultural research over the past two decades reduced the global number of people living in poverty (< $1.90 per day poverty line) by 165 million in 2015, including 110 million in rural areas and 84 million in Africa South of the Sahara. CGIAR research was, according to the analysis, associated with a reduction of about 44 million people living in poverty.

5. ***Using global agricultural, health and demographic datasets to identify the impacts of CGIAR's modern seed varieties since 1960s*** *(UC San Diego and George Washington University)*

This study uses geocoded data on births of 600,000 children in 37 countries together with newly-constructed, spatially-precise proxies of MV (modern varieties) adoption to determine the contribution of MVs to reductions in child mortality. An average 63% of cropped area was planted to MVs in 2000 in the sample countries, beginning from zero in 1960. The analysis showed that this increase in MV diffusion led to around a 3% decrease in infant mortality (from a baseline of 17%) after considering other factors likely to affect infant mortality.

The lack of high quality subnational data on MV diffusion at the required spatial and temporal resolution has hampered previous attempts to measure the effects of MVs. This analysis uses three different harvested crop area datasets to determine the relative crop mix in each location in each country: EARTHSAT data circa 2000 for 175 crops, 11 of which have modern varieties; historical EARTHSAT data for 1961-2009 covering maize, wheat and rice; and data from the Spatial Production Allocation Model which uses agricultural census data with crop suitability assessment, crop prices, and population density to derive crop area for 10 MV crops. All three datasets are available on a 5 arc-minute resolution. An MV diffusion index is constructed for each grid cell for 5-year time steps as the weighted average of crops' MV diffusion rate at that year reported at the country level by Evenson and Gollin, where the weights represent the relative share of cropped area in that grid cell devoted to that crop. Infant mortality data are obtained from the Demographic and Health Surveys which ask each mother her fertility history. The data used covers 18,382 villages in 437 administrative regions spread across 37 countries, focusing on children born from 1959 to 2001.

Infant mortality is related in a regression model to the constructed indicator of MV diffusion in the grid cell of each village, country and year together with a vector of child level controls that includes the child's sex and a quadratic function of the mother's age. The regression controls for village fixed effects and country-by-year fixed effects, including a range of binary indicators for each DHS sampling cluster, absorb all time-invariant village characteristics that could confound inference, such as climate or soils or distance to the capital city as well as any time-varying confounders at the country level, including national policy changes. The estimate of MV diffusion implicitly includes the yield-enhancing effects of input intensification that accompany the use of MVs. Therefore, effects on health should not only be considered a response to the MVs themselves, but to the wholesale adoption of more intensive and productive cropping practices stimulated by the use of MVs.

The resulting estimates suggest that the observed level of MV diffusion reduced the mortality rate of infants by 3 percentage points, which translates into 15-18 million infant deaths averted per year by the year 2000. Needless to say, these are extremely important results for CGIAR, generated by a study team entirely composed of researchers independent from the CGIAR system. These kinds of long-run and highly dispersed causal relationships are extremely difficult to detect, and yet lie at the heart of the rationale for funding international agricultural research. This study was presented at the Nairobi SIAC conference in July 2017, alongside another paper (not part of SIAC but associated with SPIA) by Gollin, Hansen and Wingender which also takes a long-run global perspective and uses modern econometric tools, finding that increased adoption of improved varieties between 1960 and 2000 reduced fertility and increased life expectancy (10 percentage points increase in HYV adoption increases life expectancy by 1.34%), and led to an increase in GDP per capita (10-15% increase for an increase in HYV area of 10 percentage points) through higher crop yields, factor adjustment and structural transformation.

6. Measuring the impact of IFPRI's research on Strengthening Food Policy through Intra-Household Analysis on the behavior of international NGOs (TANGO)

This study was designed to ascertain whether the core message of the IFPRI research program on intra-household decision-making had an impact on how international NGOs conducted their field programs. Specifically the study looked at whether international NGO program activity focused on more equitable control between men and women over allocation of resources, time, income, and program benefits had increased because of IFPRI research. The team attempted to understand: (1) the extent to which donors and international NGOs have internalized the IFPRI research results and are attempting to act on them; (2) the processes by which such results were disseminated; (3) the influence these findings had within the context of other factors influencing current approaches; and (4) whether understanding these processes could be a model for assessing the impact of policy research more broadly.

The study concludes that international NGO programs promoting more equitable control of resources by men and women increased because of IFPRI research and other factors. The authors found it unlikely that academic researchers would have undertaken such collaborative research but that the feminist network would have "continued its work and the focus on gender equality as a human right…" The element attributed to IFPRI is "the contextualization of the intra-household gender gap in specific rural settings in developing countries," which was essential for triggering action and for adapting approaches to local conditions. The methods used

in this study are inevitably somewhat subjective. Nonetheless, this study sheds light on an important and complex pathway for impact from policy-oriented research.

7. Assessing the impacts of improved cassava varieties in Nigeria *(IITA)*

This national study of cassava producing households was based on a sample of 2500 households drawn using a multi-stage clustered sampling procedure from 125 enumeration areas, representative of the four-major cassava growing areas of the country. Plot-level inputs, management and adoption data were collected for each household. Farm and plot-size data were obtained using GPS procedures. Cassava variety information was reported by farmers and confirmed using DNA analysis on samples of cassava leaves obtained from each of 7428 plots.

About 60% of farmers reported growing improved varieties of cassava whereas DNA analysis indicated that 66% were. However, about 28% of farmers reported they were growing local varieties while DNA analysis of the plant materials from their plots showed them to be using improved varieties (false negatives), and some 13% incorrectly reported growing improved varieties (false positives). Hence, based on the DNA identification procedure the "adopting" farmers were a substantially different group than those self-identified as adopters. Both groups are however likely to be selected on a number of unobservable, and the study uses an Instrumental variable (IV) analysis based on adoption rates of neighboring farmers. The IV results on self-reported improved varieties suggest increased yields of about 46% over local varieties; while IV analysis on DNA-identified improved varieties suggest yield gains of about 57%. Similar IV estimates further suggest increased per-capita food expenditure of 24%, and increases the probability of food security by 17%. These results should however be interpreted with caution as they rely on somewhat implausible exclusion restrictions. Applying an economic surplus model, the study found that the productivity gains had reduced rural poverty by 4 percentage points at a poverty line of $1.25 per person per day. The study provides important reliable estimates of large scale adoption rates on cassava in Nigeria, as well as a significant contribution to the literature in terms of its documenting the endogeneity in reporting of adoption status.

Activity 3.2. Micro-scale impact studies using experimental and quasi-experimental methods

Led by Karen Macours at the Paris School of Economics, this call was launched in mid-2014 and aimed to obtain evidence from Randomized Control Trials (RCTs) or natural experimental to broaden our understanding of specific causal pathways from technology adoption. Three studies were contracted between Nov 2014 and January 2015, and an inception workshop for the studies was held at MIT, Cambridge MA, in February 2015. In all three cases, SPIA has contributed partial funding to multi-year, costly experiments.

1. A multiple intervention approach to increasing technology adoption with a view towards scaling-up: Evidence from Mexico *(QFD Mexico, UC Berkeley, ITAM, World Bank)*

This RCT aimed to analyze whether personalized input recommendations, accounting for heterogeneity in soil characteristics across farmers or villages can improve yields. To account for potential other binding constraints, the study also addressed liquidity and additional information constraints. Over 900 farmers were successfully enrolled into the experiment, and were provided with a combination of (a) individualized plot-level soil analyses and the resulting tailored fertilizer recommendations, (b) inflexible, in-kind subsidies to purchase the tailored recommendations; and (c) high-quality agricultural extension advice. Budget considerations precluded a full factorial design so the study focused on the 4 treatments and a control.

Take-up rates of the input packages were over 75 percent in the three treatments that included subsidies to purchase recommended fertilizers. That take-up rates were comparable across these three groups suggests that neither the level of localization of the recommendations nor the restrictions on the use of the subsidy mattered. Among farmers who did not receive a subsidy for fertilizer purchases, take-up was only seven percent. In 2015, plant density was about 15 percent higher for farmers in subsidized treatments compared to farmers in the control group. The difference in the level of localization of fertilizer recommendations did not seem to affect yields to any substantial degree.

Interestingly, in the year after the intervention, farmers across treatment arms continued to use the agricultural practices learned in 2015. Farmers in the subsidized treatments reported yield increases in 2016 between 16

and 22 percent relative to the control group, although this masks considerable variation between farmers. This study suggests that providing tailored recommendations based on soil analyses can improve yields for small-scale farmers and that agro-dealers can feasibly provide such tailored recommendations to farmers. This study also documents that tailoring at an aggregate level (e.g. at the more cost-efficient village level) does not lead to significant yield losses relative to tailoring at an individual plot level. These insights suggest CGIAR research that can link soil analysis to specific input recommendations, combined with effective diffusion of such recommendations, can potentially have important and sustained impacts.

2. Drought resistance and water saving in rice production in Bangladesh *(UC Berkeley, Tufts, IRRI)*

Droughts and limited water availability are serious constraints on rice production in Bangladesh. The objective of this study is to evaluate variety BD56 – the main varietal release for drought resistance in Bangladesh. This includes (i) assessing its properties (notably in terms of yield and water saving) in the context of farmers cultivation, (ii) analyzing how farmers adjust their cropping calendar to the short duration of the variety, notably by cultivating vegetables and pulses between the main wet season and the boro rice, affecting the overall profitability through the entire cropping year, and (iii) assessing alternative ways for increasing the diffusion of the variety, by varying the selection of farmers that serve as demonstrators and how demonstrations are conducted.

The study team designed a randomized control trial where BD56 minikits were distributed to farmers in villages assigned to the treatment groups, and minikits of BD51 (the most popular variety in the region) to farmers in villages assigned to the control group. The types of farmers selected to receive the minikits differed based on the specific treatment arm that the village was assigned to in order to help understand the role of social learning and could inform how best to leverage the limited capacity available in several countries for good extension services

The study began by conducting a complete census of all 22,038 farmers in 256 study village (192 treatment, 64 control). Seeds were distributed alongside a baseline survey in July 2016, and a follow-up visit two months later that revealed that 64% of farmers had taken up the variety and had completed associated tasks for the experiment such as completing a crop calendar. Two mid-line surveys were implemented in February 2017 and in April/May 2017, and then a seed distribution effort took place in June 2017 when the all treatment villages were re-visited to test for differential rates of seed purchasing for BD56. Finally, three more surveys were to take place - on adoption (Aug 2017), investment levels by farmers (Feb 2018) and adoption (Aug 2018 – to examine diffusion through different mechanisms) before the full analysis can begin.

Preliminary results from this study suggest that the returns to BD56 are only high when farmers take advantage of its early maturation period to plant a second crop post-Aman, followed by a third Boro-rice crop. Without planting the second crop, farmers will incur a yield penalty due to BD56's short duration. However, BD56 farmers were only about 28% more likely to grow a second crop, with larger farmers twice as likely to do so. It does appear that BD56 also requires fewer irrigation days (0.5 days), which is important in these drought-prone areas where water tables are falling beyond the reach of shallow tubewells. It remains to be seen whether BD56 actually saves water in aggregate – from a temporal perspective, farmers might be risk-averse and irrigate new drought-resistance varieties lower, but not as little as they potentially could, and it remains to be seen if the number of irrigation days required declines as farmers learn.

3. Social networks to promote new agricultural technologies in Nepal *(Yale, ICIMOD)*

This randomized evaluation seeks to identify ways to improve the spread of information and increase the adoption of agricultural technologies, by varying communication structures and incentives across different treatment groups. More importantly, it also assesses how these interventions impact the welfare and productivity of farmers in a real-world setting through continuous monitoring after the initial intervention. The technological advice at the center of the evaluation is the introduction of maize inter-cropping (with either tomato, ginger, or French beans) to hill farmers in Nepal.

In total, the evaluation covered 168 wards in Nepal, with wards randomized into treatment according to different types of messenger for the extension advice – formal extension worker; lead farmer (identified as such by the community; an early adopter of technologies; literate; access to more resources than the average farmer); and peer farmer (for whom none of the restrictive conditions for lead farmers need apply). Within

each of these primary treatment arms, wards were further randomized regarding the incentives that these extension messengers receive – either no incentive, a performance-based incentive (linked to increases in adoption in the community over baseline level), or a flat incentive.

Preliminary results suggest that adoption increased by approximately 15 – 20% for all treatment groups over the control level of adoption but further analysis is pending on this study.

In June 2016, along with FERDI, SPIA organized a workshop in Clermont-Ferrand on learning from randomized control trials in agriculture. The proceedings of that meeting are available online. Earlier in the SIAC program, SPIA issued pilot project funding to three research teams – at CIMMYT / UC Davis; Paris School of Economics; and at Tufts – to help shape the development of full proposals for funding elsewhere.

Activity 3.3. Under-evaluated areas of CGIAR research

Commissioned reviews of irrigation and water management research and livestock research, highlighted areas where there are gaps in the evidence base related to impacts from CGIAR research investments. Led by SPIA member J.V. Meenakshi, a two-stage call was issued in June 2015 for impact assessment studies of under-evaluated areas of CGIAR research (irrigation & water management; livestock; agroforestry; biodiversity; policy and social science; and NRM). Proponents were asked to refer to the promising topics identified in the irrigation and water management review in particular. Four studies were eventually selected for funding in December 2015, and these are outlined below:

1. ***Forest co-management in Guinea: a multi-scale, multi-output ex-post impact analysis*** *(Virginia Tech and CIFOR)*

This study examines the impacts of the LAMIL (Landscape Management of Improved Livelihoods) Project in Guinea – a project launched in 2005 in four forest reserves in the country. The project had two planned outcomes: better governance and stronger institutions enabling more efficient and productive forestry, and to improve technical options and market linkages to increase farmers' and foresters' income. The study examines land use in a remote sensing database constructed for ten forest reserves: four forest reserves under co-management during the LAMIL project, four control reserves not under co-management during the LAMIL project, and two forest reserves previously exposed to co-management under a preceding project (ENMRA). In particular, the team identify differential changes in areas under forest reserve land type classifications in LAMIL reserve forests and non-LAMIL control reserves. Land use is also classified for 5km buffer zones adjacent to the ten forest reserves to allow for the analysis of potential land use spillovers outside of the reserves.

Complementing the examination of land-use impacts, household survey data are used to examine evidence of the sustained functioning of LAMIL reserve area Forest Management Committees (FMC) eight years after the cessation of project activities. The household survey data is also used to explore the sustained adoption of LAMIL innovations and associated impacts on household livelihoods – including indicators related to benefit flows from forest, agro-forestry, groundnut, and maize products; and household Food Consumption Scores (FCS).

The authors conclude that LAMIL project activities were associated with sustained lower rates of natural forest loss in the project and post-project periods. However, the estimates of the magnitude of retained natural forest within the LAMIL reserves are moderate – peaking at 24 square kilometer across the four forest reserves in 2014. Putting a social value on the associated retained carbon (which provides only a narrow subset of the total economic value) suggests a benefit flow of between USD 6.9 million and USD 13.8 million. The Forest Management Committees (FMCs), the main research innovation of the LAMIL project, are found to have been sustainable in the sense that they are still active and functioning and are perceived by the community as effective in three of the four forest reserves. In the fourth reserve, Nylama, the FMC is still active, but is not perceived by the community as effective in managing the forest. LAMIL appears to have increased the use of improved maize seed and the planting of tree plantations and live fencing, and there is some evidence of higher levels of food security, arising through greater diversity in the consumption of staples and fruits. However, no evidence is found of higher values of household production from maize, groundnuts, tree plantations, and live fencing – the areas of concentration of LAMIL household innovations.

This study is notable for its focus on sustained impacts eight years after the unanticipated close of the project due to widespread violence in Guinea – a rarity international development evaluation. The limited number of intervention sites severely hampers the study team in their desire to generate counterfactual evidence, but

despite this constraint, the opportunity to leverage remote sensing imagery ex post and other aspects of analysis make this study is a good example for other such cases (of which there are many in the CGIAR).

2. ***Adoption and impact of Alternate Wetting and Drying (AWD) water management for irrigated rice in the Philippines*** *(North Carolina State University, IRRI and NIA)*

The study examines the multi-dimensional impacts of AWD in the Philippines. Specifically, it represents an attempt to quantitatively analyze the micro-level economic impacts (e.g., yield, water-use, and income impact), and the environmental impacts (e.g., water saving impact at higher spatial scales and methane emission impacts) of AWD. The context for this study is a particular gravity-based irrigation system in the Philippines – the Rinconada Integrated Irrigation System (RIIS) in the Bicol Region. The study also provides some qualitative insights on the effects of AWD adoption on non-rice-producing stakeholders (e.g., hydro-electric power companies and fishpond/fish cage operators). The authors used a randomized control trial (RCT) and difference-in-differences (DID) approach.

"Turnout service area groups" (TSAGs) were randomized into a program in which farmers in treatment TSAGs were exposed (and encouraged) to adopt AWD while farmers in control TSAGs were not. 42 TSAGs were included in the study sample (21 each from treatment and control) stratified according to upstream (8 TSAGs), midstream (14 TSAGs) and downstream (20 TSAGs). The 42 TSAGs included in the study cover around 1000 ha. The 42 TSAGs were randomly chosen proportional to "size", where the size variable is the number of farmers in the TSAG. Baseline data were gathered from the treatment and control farmers prior to exposure to AWD (dry season 2016), and then follow-up data was collected after random exposure to AWD the following year (in dry season 2017). All farmers in the selected treatment TSAGs, even if they were not included in our study sample, were invited to attend the trainings. However, not all of the sample farmers attended these AWD trainings even with our strong encouragement (and even when offering to cover their transport cost to the training venue). In addition, the field water tubes (i.e., plastic PVC pipes provided by PhilRice, in this case) were personally distributed to each selected AWD treatment farmer and its use for successfully implementing AWD was explained to them. Therefore, even for those treatment farmers not able to attend the AWD trainings, information about AWD implementation was provided and encouraged.

Results from this project are based on the first analysis of the endline data using a difference-in-difference estimate, and exploiting the randomized assignment. It suggests that AWD did not have a statistically significant impact on irrigation frequency, days the main rice parcel is without water, yields, gross income or size of the main rice parcel. We conjecture that this limited impact may be because some of the control farmers already practice some form of intermittent irrigation (i.e., a limited form of AWD) where the field is NOT actually continuously flooded throughout the season.

3. ***Assessing the adoption and economic and environmental impacts of Brachiaria grass forage cultivars in Latin America focusing on the experience in Colombia*** *(CIAT, Michigan State University, Universidad de los Andes-Colombia, and CORPOICA)*

Since the early 1970s CIAT has promoted the transfer of Brachiaria grasses (native to Africa, but introduced to Latin America in the 1950s) to countries in South and Central America. This study primarily aims to estimate the current extent of adoption of Brachiaria cultivars, both in aggregate and across specific cultivars, and to identify factors that have facilitated or restricted the adoption of Brachiaria cultivars. Furthermore, the potential contribution to milk and meat productivity is estimated using an instrumental variable approach in which the presence of political turbulence or violence is used as an instrument for exogenous restrictions in Brachiaria supply to potential adopters. An extension to look at environmental impacts (such as greenhouse gas emissions) and the willingness to pay for improved forage grasses are also included in the study.

Preliminary results from the adoption study show that CIAT selected pastures represent 35% of the pastures area in Colombia. Adoption of genetically improved pastures from CIAT breeding is approximately zero, whereas the rest of the area is split between selections not carried out by CIAT and from native grasses. Only a subset of the dataset for the household survey (658 HHs, with a further 178 HHs still to come) is available,

and to date no statistically significant relationship has been found between Brachiaria grasses and income from livestock. Concerns regarding the exclusion restriction arguably limit the causal interpretation of this finding.

That said, the research is carefully conducted and provides important evidence on the relatively high levels of adoption of Brachiaria on the 8.9 MHa of pastures in Colombia. Aggregate figures for Nicaragua, Honduras, Peru and Costa Rica show much more modest shares of CIAT Brachiaria selections as a share of total pasture area in those countries. This study is the first time such a large-scale study of this important, under-evaluated research area has been conducted.

4. ***Assessing the downstream socioeconomic and land health impacts of agroforestry in Kenya*** *(ICRAF, Vi Agroforestry, and Univ. of Illinois)*

Measuring the downstream effect of agroforestry promotion on household welfare is difficult due to the long-term nature of the impact pathways and their dependence on local agro-ecological conditions. This study employs a spatial matching methodology to select a household sample prior to data collection in a long-term impact evaluation of an agroforestry project in western Kenya. Agroforestry was promoted by an NGO, Vi Agroforestry, with technical support provided by ICRAF. The extent to which the technical details of the NGO intervention can be attributed to ICRAF is addressed in the study through a rigorous qualitative approach, the results of which suggest that a large subset (but not the entirety) of the specific components/activities that comprised the intervention derived largely from ICRAF research. Specifically, while evidence of ICRAF research on improved fallows, alley cropping, and fodder shrubs influencing Vi Agroforestry's project was strong, the influence of research on tree germplasm on the program was much weaker.

The study employs a causal mediation analysis and finds that much of the increase in assets reported by households can be "explained" by the uptake of agroforestry, rather than by other components of the intervention, including sustainable land management practices. Given that ex-post matching was used to construct a counterfactual, the assumptions underlying the estimates are arguably quite strong, suggesting a clear risk of residual bias. Furthermore, the use of a long recall period for data on the pre-intervention period raises concerns about measurement error. The study team employed an array of alternative specifications and sensitivity analyses to try to mitigate these concerns. They find relatively modest effects on 'agroforestry' adoption attributable to the intervention, from which any impacts on assets is subsequently being measured.

Activity 3.4. Meta-analyses at System level: various studies

In addition to supporting individual studies, SIAC also supported meta analyses. Data on CGIAR research investments and how they have changed over time (1972-2014) were compiled to inform choices on IA portfolio. SIAC provided support to Professor Phil Pardey and his team at University of Minnesota to update work documenting global trends in investments in agricultural research and in estimating returns to those investments. The work led to several papers. One important finding is that returns to agricultural research remain very high. In Sub-Saharan Africa (SSA), Pardey et al (2016) identify 113 studies, conducted between 1975 and 2014 across 25 countries and find an average internal rate of return of 42.3% per year, with a median IRR of 35% per year.

OBJECTIVE 4: Support the development of communities of practice (CAPACITY-BUILDING)

The CGIAR will benefit from a structured attempt to support the existing capacity and some emerging collaborations on *ex post* impact assessment. Information-sharing and regular interaction are important in enabling the kinds of dialogue that can raise standards of impact assessment in the CGIAR, as well as ensuring that individuals have the skills that they need to be successful in their work. Activities towards this objective included a small grants program (activity 4.1); a targeted program of capacity-building using competitive calls for collaborations with advanced research institutes / universities (activity 4.2); conferences and workshops on impact assessment (activity 4.3); support for independently reviewing and publishing quality ratings of impact assessment studies carried out by CRPs and Centers (activity 4.4); maintenance and enhancement of the impact website (http://impact.cgiar.org) (activity 4.5).

Activity 4.1. Small grants

Four projects were funded through the small grants program (a total of US$30K) which was discontinued after 2014 owing to the heavy administrative burden for small amounts of money.

Activity 4.2. Strengthening IA capacity in the CGIAR through new partnerships

Two collaborations were funded through this activity to build capacity in the CGIAR to conduct highly credible ex post impact assessments:

1. Virginia Tech with CIFOR and CIP

Virginia Tech worked with CIP and CIFOR scientists and management on a range of activities to support the profile that impact assessment has within those institutions. A number of pilot impact assessments came out of that collaboration.

2. ICRISAT and University of Illinois

Initially limited to ICRISAT scientists, and later expanded due to demand, a workshop on advanced methods in impact assessment was held in Nairobi in September 2015 (hosted by ICRAF). The objective of the workshop was to raise the standards of impact assessment approaches and promote their application by social and bio-physical scientists at the CGIAR Centers. The agenda was based around the range of econometric tools available for impact evaluation. In general, judging by participation rates, the workshop was a success: while University of Illinois/ICRISAT were hoping to attract 25 participants, they received 52 applications (including from NARS partners), and accepted 38 applicants to attend the workshop representing 10 CGIAR Centers.

A pilot impact assessment collaboration (on conservation agriculture) between ICRISAT and Univ. of Illinois came out of this collaboration, in addition to a collaboration between ICRAF and Univ. of Illinois for a study of agroforestry/tree and fodder shrubs in Kenya (see activity 3.3 for results). The instrumental variable analysis of a 4-year panel dataset in Zimbabwe found that adoption of conservation agriculture in average-rainfall years results in no yield gains, and in some cases, results in yield losses compared to conventional practices. Conservation agriculture is, however, effective in mitigating the negative effects of deviations in rainfall: during low or high rainfall years, for specific crops (in this case, cowpea and maize), CA adoption results in higher yields. Adoption of CA is also associated with differential input costs: higher fertilizer use, but lower seed requirement.

Activity 4.3: Biennial conference on ex post impact assessment results and methods

SPIA organized workshops on specific topics in 2014 (Minnesota – Measuring poverty impacts of CGIAR research) and 2016 (Boston – Innovative methods for measuring adoption of agricultural technologies). The conference in Nairobi in July 2017 was the culmination of several years' work across the SIAC program and was attended by 180 participants.

Activity 4.4. Enhancing quality and rigor: Introducing a Star Rating System for IA studies

While the Editorial Manager (a typical journal management platform) was customized and launched by SPIA and Allen Press as of February 2015, the activity has not progressed further, even if discussions on approach to quality rating and criteria to be used have continued (criteria have been revised further). When we trialed the system in 2015, requesting scientists to voluntarily submit their work for quality rating, we did not have any uptake. We are considering whether to use the system in a more top-down process in which we identify specific studies that need external review.

Activity 4.5. CGIAR Impact Website

Rather than persisting with a stand-alone website, we have now migrated all the content from http://impact.cgiar.org over on to www.ispc.cgiar.org. This should help drive traffic to the impact assessment pages (http://ispc.cgiar.org/workstreams/impact-assessment).

Annex

LIST OF SIAC FUNDED STUDIES

No.	**Grantee** / Sub-grantee - Partner	SIAC activity	Project start	Project end	Project extensions / terminations	Status
1	**Michigan State University**	OBJ.1/2	01.03.13	31.12.15	31.10.16	completed
2	**Michigan State University**	OBJ.1/2	01.02.15	31.05.16	31.10.16 / 31.12.16	completed
3	**VIRG TECH** / CIFOR	OBJ. 4.2	01.01.14	31.12.15	30.11.16	completed
4	**ICRISAT** / Univ Illinois	OBJ. 4.2	01.10.14	30.09.16	31.12.16	completed
5	**UIUC**	OBJ. 4.2	01.01.17	15.05.17	-	completed
6	**CIMMYT**	OBJ.3.2	01.01.14	01.03.14	-	completed
7	**TUFTS**	OBJ.3.2	01.03.14	01.03.15	28.02.16	completed
8	**Paris School of Economics**	OBJ.3.2	01.01.14	01.03.16	-	completed
9	**Innovations for Poverty Action (IPA)** / MIT-JPAL, MIT Sloan	OBJ.3.0	15.07.14	31.12.15	31.12.16	completed
10	**ILRI** / Emory Univ	OBJ.3.0	01.07.14	30.06.16	30.09.16	completed
11	**CIMMYT** / Univ Goettingen, Univ Geneva	OBJ.3.0	01.10.14	30.06.16	-	completed
12	**CIAT** / Virg Tech	OBJ.3.0	15.11.14	14.11.16	14.05.17	completed
13	**COLUMBIA** / George Washington Univ	OBJ.3.0	01.03.15	31.12.16	21.11.16	terminated
14	**Johns Hopkins University**	OBJ.3.0	01.01.17	31.05.17	-	completed
15	**Qué Funciona para el Desarrollo A.C** / Berkeley, J-PAL, WorldBank	OBJ.3.2	01.12.14	31.10.16	-	completed
16	**Qué Funciona para el Desarrollo A.C** / Berkeley, J-PAL, WorldBank	OBJ.3.2	06.12.16	31.05.17	-	completed
17	**BERKELEY** / Tufts, IRRI	OBJ.3.2	01.05.15	31.12.16	31.05.17	completed
18	**YALE** / J-PAL, ICIMOD, Dep. of Agric. Nepal	OBJ.3.2	01.05.15	31.12.16	30.04.17	completed
19	**ICARDA** / Virg Tech	OBJ.3.1	01.07.15	31.12.16	31.03.17	completed
20	**CIP** / Virg Tech	OBJ.3.1	01.06.15	30.11.16	29.02.17	completed
21	**IITA**	OBJ.3.1	01.06.15	31.05.17	-	completed
22	**UC SAN DIEGO** / Ram Fishman	OBJ.3.1	01.07.15	31.12.16	31.05.17 / 30.06.17	completed
23	**IFPRI** / WorldBank	OBJ.3.1	01.08.15	31.12.16	31.03.17	completed
24	**WORLDFISH** / Central Luzon Univ Philippines, Phiippines Bureau of Fisher	OBJ.3.1	01.08.15	31.12.16	31.05.17	completed
25	**TANGO International** / IFPRI, Westat	OBJ.3.1	15.10.15	31.08.16	31.03.17 / 31.05.17	completed
26	**CIAT** / Michigan State Univ, Univ Andes	OBJ.3.3	15.02.16	30.06.17	-	completed
27	**VIRG TECH** / CIFOR	OBJ.3.3	01.03.16	30.06.17	-	completed
28	**IRRI** / North Carolina State Univ / PhilRice	OBJ.3.3	15.02.16	30.06.17	-	completed
29	**ICRAF** / Univ Illinois	OBJ.3.3	01.02.16	30.06.17	-	completed
30	**University of Minnesota**	OBJ.3.5	15.12.15	15.11.16	-	completed
31	**Institute for Financial Management and Research** / CIMMYT	OBJ.2.2	15.05.16	30.06.17	-	completed
32	**Norwegian University of Life Sciences** / Monica Fisher	OBJ.2.2	15.05.16	30.06.17	-	completed
33	**CIMMYT** / ICRISAT	OBJ.2.2	15.05.16	31.05.17	-	completed
34	**Nong Lam University Vietnam** / UC Santa Cruz	OBJ.2.2	15.07.16	30.06.17	-	completed
35	**IFPRI** / IITA / GeoPoll	OBJ.2.2	15.07.16	30.06.17	-	completed
36	**ICRAF** / Penn State, Univ Zambia	OBJ.2.2	01.06.16	30.06.17	-	completed
37	**CIMMYT**	OBJ.2.2	01.06.16	30.06.17	-	completed
38	**ICRISAT** / Univ Illinois / CIMMYT	OBJ.2.2	01.07.16	30.06.17	-	completed
39	**FAO EPIC**	OBJ.2.2	15.07.16	30.06.17	-	completed

(sorted in cronological order by date of the call and grouped, in colour scales, by the SIAC objective)