Photo © Dominic Chavez/World Bank

# Impact of CGIAR's Agricultural Research for Development: Findings and Lessons from the SIAC Program

CGIAR

Independent Science and Partnership Council

Standing Panel on Impact Assessment (SPIA)

**ABOUT SPIA**

The Standing Panel on Impact Assessment ('SPIA') is an external, impartial standing panel of experts in impact assessment that advises the CGIAR system. SPIA's mandate is to expand and deepen evidence of impact of CGIAR research investments on CGIAR Strategic Results Framework (SRF) outcomes and associated Sustainable Development Goals (SDGs), and to support CGIAR's strong commitment to embed a culture of impact assessment into the System.




**Acknowledgements:**

As a summary of the impact assessment work undertaken under SIAC, this report reflects the work of many people who contributed to the design and implementation of the SIAC program—including external reviewers and members of the Project Steering Committee (PSC), and to the 29 individual studies that the SIAC program supported. The report was compiled and written by Lakshmi Krishnan and Nancy Johnson from ISPC/SPIA Secretariat and SPIA Chair Karen Macours. JV Meenakshi, Doug Gollin, Robert W Herdt, James Stevenson, Mitch Renkow, and Jeff Alwang provided valuable feedback. Ira Vater compiled the IA budgets data, and provided support to overall administrative, financial, and management activities throughout the SIAC program. Finally, SPIA would like to specially acknowledge and thank Timothy Kelley (Senior Agricultural Research Officer, ISPC/SPIA Secretariat, 1999-2016) for his substantial contributions to the conceptual design, development, and excellent stewardship of the SIAC program.

**Funding Acknowledgements**

This research was supported by ISPC-SPIA under the grant 'Strengthening Impact Assessment in the CGIAR (SIAC).'

**Standing Panel on Impact Assessment (SPIA)**

https://ispc.cgiar.org/workstreams/impact-assessment

Cover image: Photo © Dominic Chavez/World Bank

Design and layout: Macaroni Bros

Editor: Heidi Fritschel

# Impact of CGIAR's Agricultural Research for Development:

# Findings and Lessons from the SIAC Program

February 2019

# TABLE OF CONTENTS

# EXECUTIVE SUMMARY

This report characterizes and synthesizes the findings from 25 impact assessments funded under SPIA's Strengthening Impact Assessment in CGIAR (SIAC) program between 2013 and 2017 (See Table 2 for full list). The studies examine the impacts of a wide range of innovations from three core areas of CGIAR research—crop improvement, production systems management, and policy. Studies used a range of methods, from RCTs to qualitative approaches, depending on the specific research question being addressed.

The majority of the studies looked at impacts of improved varieties and at outcomes related to poverty and food security. That said, the range of crops covered went well beyond the major cereals, and SIAC studies also add to the evidence base on nutrition and health outcomes and, to a more limited extent, on environmental outcomes of agricultural research. This high-quality evidence of adoption and impact at scale, rather than for a small number of villages or pilot sites, is a significant addition to the evidence base.

Five studies generated new evidence on **technology adoption at scale**—for cassava, potato, lentil, beans, fish, and tropical forages. One study documented the adoption of agroforestry innovations and other natural resource management practices in the context of a long-term development program.

**Poverty (SLO 1)**. Adoption of improved cassava varieties from IITA was associated with a reduction in poverty in Nigeria. Six nationally or subnationally representative studies report positive impacts of improved varieties and production system innovations on intermediate outcomes such as income, consumption, resilience or assets. At the global scale, productivity increases resulting from the adoption of modern varieties associated with CGIAR contributed to a 139 percentage-point increase in overall GDP per capita between 1960 and 2000 in a sample of 84 countries, in part by allowing for structural transformation.

**Food security and nutrition for health (SLO 2**). Two global studies provide important new evidence on the nutrition, health, and demographic impacts of the Green Revolution. Farmers' adoption of modern varieties between 1960 and 2000 was associated with a significant long-run reduction in infant mortality (3-5 million infant deaths averted per year by the year 2000), and contributed to reductions in both overall mortality and fertility, with a net negative effect on population growth.

Studies of more recent impacts suggest that changing crop varieties and production practices have some potential to contribute to improved food security and nutrition status, particularly within producer households. The evidence is not conclusive, however, and the evidence base needs to be broadened to include the nutrition and health impacts of market and governance interventions (Bulte, 2018).

**Natural resources and ecosystem services (SLO 3)**. Two studies that looked at development programs informed by CGIAR research document positive environmental outcomes—in one case a reduction in the rate of loss of natural forest in Guinea, and in another an increase in tree cover in Kenya. Many other studies with potential natural resource or environmental implications did not measure such outcomes, even when the studies hypothesized that they existed. This suggests that the SIAC portfolio might be underestimating impacts—positive and negative—of CGIAR research on SLO 3.

Some key methodological lessons from the SIAC portfolio include:

- **The portfolio illustrates improved rigor and innovative use of tools for quasi-experimental methods and modeling approaches, but more can be done**. These approaches are appropriate for looking at long-term impacts that go beyond plot- or farm-level productivity, and they can speak to the kinds of impacts many stakeholders are interested in. Improving the quality of data and methods and increasing the number of studies—so that results of studies of similar innovations conducted in different contexts using different methods can be compared and synthesized—will help strengthen the evidence base. Investing in regular collection of data, at representative scales, on adoption of CGIAR innovations and related outcomes will expand the number and the quality of studies that can be conducted.

- I**n most cases, little attention was paid to synergies and trade-offs among outcomes**. While some SIAC studies, in particular natural resource management studies, were designed to examine the trade-offs that households face, most did not do so systematically. A few of those that did found evidence of potential trade-offs, especially related to risk management. This suggests that impact pathways should be screened more systematically at an early stage to identify potential trade-offs and address them during research or impact evaluation.

- **Very little attention was devoted to gender and wider socioeconomic heterogeneity**. With a few exceptions, the SIAC studies did not consider gender or other relevant socioeconomic dimensions. Even disaggregation of outcomes by sex or other relevant socioeconomic dimensions was done in only a few studies. More attention should be paid to ensuring that heterogeneity is well addressed in future CGIAR impact assessments, building on the work that has been done to integrate gender into research program theories of change.

Some of the main findings in relation to data and measurement are:

- **Use of DNA fingerprinting to identify crop and fish varieties was demonstrated to be feasible at scale. In many cases, it was also shown to be necessary to accurately estimate adoption and impact**. Even where it may be feasible to get accurate aggregate modern variety adoption estimates from, for example, self-reported data, observed type I and type II errors in self-reported varietal data illustrate why fingerprinted data are better for analyzing correlates of adoption. However, practical issues like sampling, preserving tissue samples from deterioration, and generating reference libraries all require careful attention (particularly in developing-country field conditions) to ensure that DNA fingerprinting is carried out correctly.

- **The use of remotely sensed data and geospatial tools to document adoption and impact shows much promise**. It can reduce the cost of data collection, improve the credibility of studies where a baseline is absent, and allow association of datasets with other data that are not available at required spatial scales (for example, when rain gauges are not available at the village level) or are costly to collect (for example, data on soil fertility).

- Finally, for all innovations, there is **a need for better measures of productivity and broader measures of potential benefits and costs beyond yields**. While many studies documented adoption, less attention was paid to careful measurement of land and labor productivity, income, or profitability. This information is essential to understand how farm households benefit (or do not benefit) from particular innovations, which has implications for predicting adoption and for convincingly linking it to poverty reduction or other outcomes.

# FOREWORD

In an era where public sector funded agricultural research is being called upon to address problems of increasing complexity, whilst budgets are tightened and closely scrutinized, the ability to track impacts across a range of outcomes has become a high priority. CGIAR researchers face considerable pressure to demonstrate convincingly the value of their work in achieving sustainable development outcomes, including the elimination of poverty and food insecurity, improving nutrition, and natural resource management. Rigorous *ex post* impact assessments are an essential tool for developing credible and relevant evidence of the contribution of agricultural research to sustainable development. This report represents a significant step forward in understanding the impacts of CGIAR research activities, providing a summary of findings from an unprecedented set of rigorous assessments. It also gives important insights into how impact assessment methods for agricultural research can, and should be improved. As such, it is of interest to investors in agricultural research, as well as impact assessment practitioners in this field.

The 25 new impact assessments synthesized in this report were commissioned by the CGIAR Standing Panel on Impact Assessment (SPIA) between 2013 and 2017 under the Strengthening Impact Assessment in CGIAR (SIAC) program. The SIAC portfolio of impact assessments includes both older and relatively recent investments, covering a wide range of CGIAR research activities, from crop germplasm and production systems to gender, forest management, and policies and institutions. The SIAC program encouraged the experimentation and use of a wide range of methodologies (traditional *ex post* as well as impact evaluations).

One of the important findings from these studies is the positive and at scale impacts of agricultural research in contributing to sustainable development. They also indicate that the impacts unfold over extended periods—even decades. For instance, farmers' adoption of modern varieties of crops between 1960 and 2000 was found to be associated with reductions an infant mortality—at least about 3-5 million infants avoided premature death across 36 countries. Five large-scale adoption studies across six country contexts provide new evidence of 3.61 million farm households and 4.11 hectares of crop or grassland area under technologies related to CGIAR research. However, the findings also indicate that the introduction and adoption of new technologies does not automatically translate into welfare benefits, as evidenced by the results from two randomized control trials of the impacts of early maturing varieties.

In recent years, greater emphasis has been given to linking research to development outcomes, and CGIAR research informing development projects has been an important means of "scaling up and out" research to development impacts. Two of the SIAC studies were devoted to assessing impacts of this type of intervention, which indicated positive impacts in fairly narrowly defined dimensions.  They also indicated the degree of complexity both in achieving research to development impact, and in measuring it.

The issue of assessing impacts in complex programs is particularly evident in the studies focusing on NRM interventions, where institutional and policy changes are particularly important to achieving impact at scale. Impact assessments narrowly focused on field level outcomes cannot capture the impacts research may have had on policy and institutional change and thus may not be giving us the full picture of the research impacts.

One of the most striking results emerging from this analysis is the weak capacity to assess a broad range of development impacts—including changes in ecosystem services and their implications for agricultural productivity, nutritional outcomes or gender impacts. This is clearly an area where greater attention is needed in future CGIAR impact assessment work. Likewise, the report indicates a weak capacity for examining synergies and tradeoffs that can arise between environmental, social and economic outcomes resulting from adoption of technologies and innovations and this is a critical need if CGIAR research is to effectively contribute to meeting the 2030 Sustainable Development Agenda.

Overall, the report represents a milestone in broadening our understanding of the impacts of past agricultural research activities, as well a beacon to guide ongoing assessment work to provide a broader and more accurate picture of the impact of CGIAR work.

**Leslie Lipper**
**Executive Director, ISPC Secretariat**

# 1 INTRODUCTION

CGIAR has long been a leader in estimating the impacts of and the rates of return to investment in international agricultural research (Rao, Hurley, and Pardey, 2016). In recent years, though, assessing and documenting evidence of development impacts has become more challenging owing to the broadening of the research and development agenda and increased recognition on the part of donors and other stakeholders of the need for higher standards of evidence for impact assessment (Baylis et al., 2016; Cameron, Mishra, and Brown, 2016; Duflo and Kremer, 2003; Savedoff, Levine, and Birdsall, 2006). The Strengthening Impact Assessment in the CGIAR System (SIAC) program was an investment by CGIAR donors between 2013 and 2017 to begin to fill gaps in data, evidence, and internal capacity for impact assessment in light of this changing context (ISPC, 2017). This report focuses on SIAC's contribution to the evidence base on impacts of CGIAR research, summarizing and synthesizing the findings of 25 impact assessment studies (resulting in 26 papers) that received financial and/or technical support from CGIAR's Standing Panel on Impact Assessment (SPIA) through SIAC. It complements recent systematic reviews (Garbero, Marion, and Brailovskaya, 2016; Pray, Masters, and Ayoub, 2017) that have examined the impacts of agricultural research for development (AR4D) investments on productivity, poverty, and nutrition. The report also reflects on lessons learned in SIAC about how to design and implement an impact assessment portfolio for CGIAR—that is, at a system level.

Linking CGIAR research to improvements in development outcomes, especially the three goals of reducing poverty, improving food and nutrition security for health, and improving natural resources and ecosystem services[1] at scale, is not an easy task. The causal chains through which advances in agricultural research influence development outcomes are nonlinear and complex, and they frequently involve long time lags, making impacts inherently difficult to measure (Gollin, Probst, and Brower, 2018). Multiple sources of evidence from studies using different methodological approaches—micro and macro, quantitative and qualitative—are needed to piece together a convincing *ex post* case for contributions from research, particularly at the institutional or programmatic level. Any given study will provide useful evidence on some impact questions while being unable to address others in a convincing manner, even with the best available data and methods. Recognizing this, SPIA's goal for SIAC was to assemble a portfolio of studies that would improve on the existing evidence base regarding the Strategy and Results Framework as a whole, and CGIAR's aspirational targets as outlined in the SRF 2016–2030.

[1] Reducing poverty, improving food and nutrition security for health, and improving natural resources and ecosystem services are the System Level Outcomes (SLOs) defined in CGIAR's Strategy and Results Framework (SRF) 2016–2030 (CGIAR 2016).

The report is organized as follows. Section 2 describes the portfolio of SIAC studies and how it was constructed to respond to the challenges of assessing impact at a system level for an organization like CGIAR. Section 3 describes the studies by area of research, and section 4 presents the main findings on development outcomes and impacts (technology adoption, poverty, food and nutrition security, and environment) of past research. With a focus on the learning agenda, section 5 reports results from studies that used experimental methods both to rigorously estimate effects of innovations on outcomes and to better understand key links in casual pathways. Section 6 concludes with reflections on SIAC evidence, and the main strategic and methodological lessons.

# 2 STRUCTURING A PORTFOLIO TO ASSESS SYSTEM-LEVEL IMPACT

SPIA's mandate is to provide evidence on system-level impact, complementing the impact evaluation and assessment work taking place in programs and projects[2] of the 15 CGIAR research centers (Kelley, Ryan, and Gregersen, 2008). SIAC thus aimed to broaden and deepen understanding of the nature and extent of the impacts of CGIAR research. In defining the portfolio for SIAC, several factors were considered:

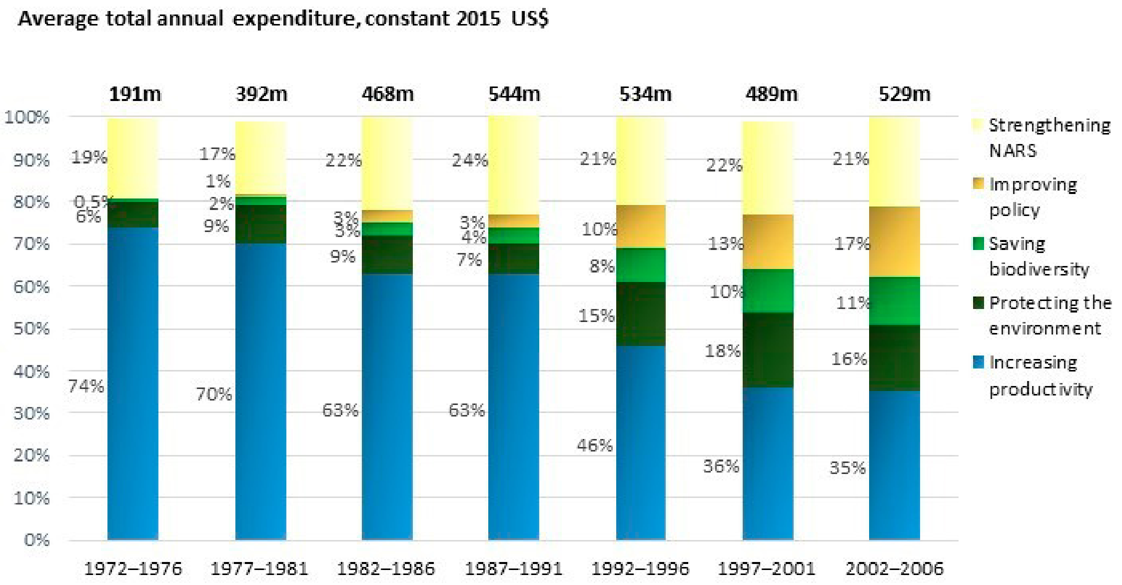
- **System-level investment patterns and trends**

An obvious starting point for designing an impact assessment portfolio is to look at past research investments. Tracking research investments is important both for understanding investment patterns and for identifying specific areas of research for impact assessments. Aggregating data from centers to obtain investment levels and trends at CGIAR system level is surprisingly challenging and is possible only through 2006 (Elven and Krishnan, 2018; Pingali and Kelley, 2007). Figure 1 shows investment by major research objective (termed "research area" in CGIAR records) for 1972-2006. Research on increasing productivity was the largest area in absolute and relative terms in 2006, though the portfolio had been diversifying since the early 1990s. Investments in natural resource management[3] (mapped as "protecting the environment") and in "improving policy" have increased since the early days of CGIAR. Research on both germplasm improvement and production systems management was initially classified under "increasing productivity," but that changed in the 1990s when some production systems work appears to have been characterized as "protecting the environment" (Elven and Krishnan, 2018).

[2] CGIAR's Independent Evaluation Arrangement (IEA) is compiling and analyzing an inventory of *ex post* impact assessments conducted by CGIAR centers since 2010. According to a preliminary assessment, 16 percent of the 198 studies are in some way linked to SPIA.

[3] The research area "saving biodiversity" refers to germplasm collection, characterization, and evaluation—that is, genebank-related activities, not *in situ* conservation or management of natural resources.

**Figure 1.** CGIAR expenditure by research area, 1972–2006




**Source:** Elven and Krishnan, 2018, p. 3.

**Note:** *NARS = national agricultural research systems.*

### Current impact assessment evidence base

The bulk of CGIAR impact evidence is on the productivity and economic impacts of crop germplasm improvement, especially yields of the major cereals (rice, wheat, maize) (Kelley, Ryan, and Gregersen, 2008; Renkow and Byerlee, 2010). As CGIAR's mandate expanded from productivity to poverty and sustainable resource management in the 1980s and 1990s (see for instance Gryseels et al., 1992), SPIA made efforts to document impacts on these additional outcomes in the past (Kerr and Kolavalli, 1999; Meinzen-Dick et al., 2004; Stevenson et al., 2013; Waibel and Zilberman, 2007).

The 2012 CGIAR reform introduced nutrition and health as explicit outcomes at the system level. The contributions that agriculture makes to nutrition and health through improved food security and economic growth have long been recognized (Evenson and Rosegrant, 2003). The new Strategy and Results Framework (SRF) highlights additional pathways) through which agriculture could contribute more directly to improved nutrition and health outcomes (see for instance Webb, 2013), most notably through improving diet quality and food safety. This new focus of the SRF hence called for widening the evidence base.

### Balance between accountability and learning

Impact assessment (IA) has both accountability and learning functions. In accountability-oriented studies, the main questions are whether a specific research investment contributed to the intended outcomes and impacts and whether the magnitude or value of the benefits helps justify the investments in the overall research system. These effects can be observed only long after research is completed and research outputs have been widely taken up and used. As a result, impact assessments that address these higher-level outcomes tend to focus on older outputs that are considered "successful." While cherry picking such successes in designing an accountability-oriented IA portfolio may make sense—in research, a few large successes generally justify the whole portfolio (Raitzer and Kelley, 2008)—the value of results from such studies in informing future research priorities and investments is often limited. This is because the context in which future research outputs will be used is likely very different from the context in which past outputs contributed to development impacts.

Learning-oriented impact assessment is focused on understanding how and why research contributes (or not) to impacts, for whom, and under what circumstances. Results from learning-oriented studies can more easily inform scaling-up strategies for research dissemination and feed back into the research agenda-setting processes. Such studies are often conducted soon after research is completed[4] and before it is widely scaled—or if an output fails to scale as expected. Experimental methods are often well suited to answer these types of questions.

- **Assembling the SIAC portfolio**

To assemble a portfolio of studies, SIAC issued four competitive calls for proposals (Table 1). One specifically addressed learning (short-term, micro-scale) and another accountability (*ex post* IAs of scaled-up or older research innovations). One call targeted the agriculture-nutrition evidence gap. The last call to be issued focused on "underevaluated areas" and was informed by a series of scoping studies (Jutzi and Rich, 2016; Merrey, 2015) as well as work commissioned on policy-oriented research (Renkow, 2018). The total funding allocated across studies was US$3.7 million (Table 1). In addition to the budget allocated directly to the calls, an average of US$55,000 was budgeted to provide technical oversight, commission external reviews, and conduct inception or final workshops with study leaders.

The modalities of each call varied slightly, but in general they involved external reviews of proposals as well as technical reviews and support from SPIA to improve studies at the conceptual stage. Proposals were scored and selected on technical merit, with due consideration given to whether the proposal filled an important evidence gap. Partnerships between CGIAR researchers and external impact assessment specialists were strongly encouraged, with SPIA theorizing that such partnerships could enhance the quality and relevance of individual studies—a gap identified in the 2009 Social Science Stripe Review (CGIAR Science Council, 2009).

**Table 1.** SIAC impact assessment calls for proposals[5]

| Competitive call window | SPIA budget (US$) | Share of total study budget from SPIA[a] | # of studies funded |
|---|---|---|---|
| 1. From agricultural research to impact on nutrition and health outcomes, July 2013 [Link] | US$679,995 | 35% | 5 (6 papers) |
| 2. Short-term micro-scale studies using experimental and quasi-experimental methods, April 2014 [Link] | US$1,022,364 | 48% | 5 |
| 3. *Ex post* IAs of scaled or older research *innovations*, July 2014 [Link] | US$1,251,548 | 52% | 7 |
| 4. *Ex post* IAs of underevaluated research areas, June 2015 [Link] | US$739,684 | 71% | 4 |

[a] Sources of co-funding varied across the individual studies and were among the criteria considered by reviewers in the process of selection.

[4] Ideally, real-world efficacy and effectiveness studies would also be conducted during the research process itself as innovations are being tested and developed.

[5] The source for study budgets in these tables is the initial letters of agreement. It is possible that these budgets and sources of funding changed over the course of implementation – as IA researchers sought additional (non-SIAC) funding.

SIAC also issued a separate competitive call to develop impact assessment capacity through partnerships between CGIAR research centers and IA experts external to CGIAR, typically in universities. This window funded a wide range of activities—including methodological workshops, in-person feedback and discussions on IA portfolio prioritization, postdoctoral positions—and resulted in four additional impact assessments that are included in this report. Altogether, this synthesis covers 25 impact assessments (resulting in 26 papers) that received support from SIAC to assess impacts of CGIAR research on SRF outcomes (see Table 2).

It is worth noting that the total budget from SPIA for these studies was just under half of the full cost of the studies—ranging from 35 percent in the "nutrition and health" call to 71 percent in "undervaluated areas" call (Table 1). This underscores the fact that most studies were co-funded, often with SIAC funds leveraging additional data collection and/or analysis in the context of already-funded studies. Of the 25 studies, CGIAR centers were involved in conducting all but eight, half of which were randomized controlled trials (RCTs) requiring specialized expertise.

# 3 SIAC PORTFOLIO BY RESEARCH AREA

This section describes the distribution of studies across three research areas[6]: germplasm improvement and breeding (11 studies); production systems management (nine studies, resulting in ten papers); and improving policy (five studies) (see Table 2, column 3). None of the studies focused primarily on protecting the environment; however, two of the policy studies were directly related to forest management and conservation, and a number of studies measured or hypothesized environmental outcomes. No single study focused on strengthening national agricultural research systems (NARS); however, capacity strengthening (of, for instance, farmers and extension agents) was an outcome in many impact pathways, and aspects of capacity were measured in several studies, especially the learning-oriented studies.

## 3.1 GERMPLASM IMPROVEMENT AND BREEDING

Of the 11 studies on germplasm improvement and breeding (Table 2), six focused on a single crop, one on a forage grass, and one on a strain of fish. Three global studies relied primarily on existing datasets from multiple countries and CGIAR commodities. Another study looked at improved tree species—however, since it was assessed as part of a broader agroforestry systems intervention, the study is classified under production systems management (see section 3.2). Of the single-crop studies, four looked at noncereals—beans, cassava, lentils, and potato—for which few rigorous adoption and impact studies at scale were previously available. Methodologically, in addition to simply asking farmers the names of the varieties they grew on their plots (as part of household surveys), the bean, cassava, lentil, potato, and fish study teams collected seed and/or tissue samples to identify varieties using DNA fingerprinting (Box 1). Most genetic improvement studies assessed the impact of research initiated in the 1980s and 1990s (Table 2). Most were large scale in nature, in most cases regional or national.

The three studies that assessed more recent research efforts focused on understanding the adoption and impact of novel traits (micronutrient content, early maturity) and/or outcomes (nutrition) in experimental or development program settings. The results of these types of studies are an important complement to the results of large-scale studies, providing rigorous evidence on the magnitude and distribution of impacts on target populations and identifying factors that could enhance or constrain uptake at scale.

[6] These categories, from Elven and Krishnan (2018), are a more disaggregated version of those in Figure 1.

## BOX 1. MEASURING VARIETAL ADOPTION USING DNA FINGERPRINTING

Data on adoption of crop varieties by farm households are costly to collect. Traditionally, data collection has relied on farmer surveys, and when those have been unavailable or too costly to implement, expert opinion elicitation has been used, especially at the national or regional scale. DNA fingerprinting has been used in agricultural research for many years, but the recent dramatic decline in the cost per sample has made it feasible to integrate this method into farmer surveys, with samples collected directly from farmers' fields, potentially even at a large, statistically representative scale. Under SIAC, SPIA commissioned a series of pilots in different contexts to compare data derived from DNA fingerprinting with expert opinion data as well as with a range of survey-based methods for eliciting adoption information. This work is part of a growing literature on measurement experiments in agriculture (Beegle, Carletto, and Himelein, 2012; Carletto et al., 2016; Kilic and Sohnesen, 2015). According to the SIAC pilots, expert elicitation typically resulted in an overestimation of the adoption of improved varieties, while estimates based on farmers' own identification were closer to DNA fingerprinting results. Even so, important measurement error was also identified (more on this below).

While some of the early SIAC experiments were small-scale pilots designed to uncover the extent of measurement errors in traditional methods, others were implemented at scale as part of adoption and impact studies. Four commodity studies all used DNA fingerprinting tools to identify varieties from samples collected from farmers' fields or fish nurseries. Doing so greatly increased confidence in the variety identification data but at an increased cost, raising the question of when it is optimal to use DNA fingerprinting approaches.

In the study by Wossen et al. (2017), 28 percent of cassava farmers in Nigeria reported that they grew local varieties whereas the DNA data showed that they in fact grew improved varieties (false negatives). Correspondingly, about 13 percent reported that they grew improved varieties whereas the DNA showed that they grew local varieties (false positives). By contrast, in Bangladesh, where lentils had been introduced relatively recently, there was approximately a 90 percent correspondence between the DNA-based varietal identification and farmer self-reported data (ICARDA, 2017). In China, 97 percent of the 141 samples reported to be C88 potato were confirmed by DNA analysis to be so, suggesting low levels of false positives in the self-reported data. However, no DNA testing was done on potatoes grown by the other 475 households, so the picture is incomplete—that is, no data are available on false negatives (CIP, 2017). A study of Genetically Improved Farmed Tilapia (GIFT) provides detailed adoption data validated by DNA (Benzie and Lind, 2017) showing GIFT or GIFT-derived strains of tilapia were the most commonly found strains in sampled nurseries in both Bangladesh and the Philippines. The International Center for Tropical Agriculture (CIAT) intended to use DNA fingerprinting to identify bean varieties in a nationally representative adoption study of high-iron beans, but technical problems resulted in spoilage of samples before lab analysis. Studies of cassava in Nigeria (Wossen et al. 2017) and lentils in Bangladesh (International Center for Agricultural Research in the Dry Areas [ICARDA], 2017) found that incorrect classifications influenced findings on the correlates of adoption and, in the cassava case, could also lead to wrong inferences regarding productivity impacts.

If the objective is to characterize adoption or measure impact, valid data on both the genetics and the production inputs are needed. The SIAC results on DNA fingerprinting, which are being summarized in a scientific publication and a practitioners' guide, show that in some cases DNA fingerprinting is essential to accurate estimation. However, determining if the estimation of adoption for a specific crop-by-country combination requires DNA fingerprinting (or not) is a challenging task. Bangladesh and China had most lentil or potato seeds moving through formal systems, a relatively small number of modern varieties in use on farmers' fields, and very low use of landraces (none for potatoes in China, less than one percent for lentils in Bangladesh). But, even when it may possible to use such contextual information to make an informed decision to prioritize resources for DNA fingerprinting in order to derive adoption estimates, one would still need to validate such hypotheses (on the degree of confidence in farmer self-reporting) with limited testing. When fingerprinting is used, practical issues like sampling, preserving tissue samples from deterioration, and generating genetic signatures of standard varieties (reference libraries) for comparison with samples all require careful attention, particularly in developing-country conditions, in order to ensure that DNA fingerprinting results accurately reflect the crop varieties present in farmers' fields.

**Sources:** Herdt (2018); Stevenson, Gollin, and Macours (2018).

## 3.2 PRODUCTION SYSTEMS MANAGEMENT

Ten SIAC studies examined CGIAR-related innovations in production and management systems—three on intercropping of maize with legumes or vegetables; two each on water management, soil fertility management, and conservation agriculture; and one on agroforestry (Table 2). In general, the research assessed in these studies is more recent than in the genetic improvement studies. Most studies looked at the outputs of research undertaken in the 1990s and 2000s.

Four studies sought to document impact at scale—though what is considered "scale" for these types of innovations is, for various reasons, generally smaller than for improved varieties. The rest focus on understanding the benefits of the innovations, how they are distributed across different types of farms and farmers (heterogeneity), and what the implications are for dissemination strategies. These learning-oriented studies, most of which were funded under the RCT call, respond to a growing recognition that many technologies—especially complex technologies such as production system innovations[7]—are often not as widely adopted as expected in the absence of dissemination programs (Stevenson and Vlek, 2018). Whether this low adoption is due to deficiencies inherent in the innovations themselves or to the challenges associated with effective dissemination (including targeting) is not conclusively known, and is a key question these studies sought to answer.

In terms of outcomes assessed, all ten studies focused on yield and associated economic benefits. Four sought to measure food security and nutrition-related outcomes. Despite the fact that the research had a clear "systems" focus, only two studies—on agroforestry and alternate wetting and drying (AWD)—measured the private or public natural resource–related benefits (such as soil fertility or water savings) associated with these innovations.

## 3.3 IMPROVING POLICY

Influencing programs and policies is both an impact pathway (Box 2) and an area of research in CGIAR. Five studies evaluated how social science research results influenced policies and programs. Two studies were about improving the implementation of forestry policies (governance and management), two focused on improved value chains (dairy, forest products), and one looked at how a better understanding of gender relations in agricultural households influenced non-governmental organization (NGO) programming. As in the case of production systems management, the research assessed in these studies is relatively recent—spanning the late 1990s to the 2000s.

As part of SIAC, a database of 94 plausible claims of policy influence by CGIAR was compiled covering the period 2006–2014 (Renkow, 2018). Four of the five impact assessments supported under SIAC—two on forestry, one on dairy hubs, and one on gender—are attempts to follow up on those claims to more rigorously document outcomes and impacts.

[7] In the classification of CGIAR expenditure data, production system research was initially considered a subset of "Improving productivity" rather than "Protecting the environment." The reality, especially since the 1990s, is that it addresses both the productivity and the sustainability/natural resource management (NRM) agendas and is less tied to the productivity of the mandate crop (Waibel and Zilberman, 2007).

## BOX 2. ASSESSING THE EFFECTIVENESS AND IMPACT OF RESEARCH IN THE CONTEXT OF AGRICULTURAL DEVELOPMENT PROGRAMS

Development programs implemented by multilateral institutions, NGOs, or governments are an important pathway through which CGIAR research reaches end users. For some types of innovations with significant public good characteristics, they may be the primary mechanism. Development programs can also be valuable opportunities to learn about the short- or medium-term effectiveness and impacts of research outputs. Conducting impact assessments in the context of development programs requires not only documenting the link between the research and the program—which can take several forms—but also establishing a "nonresearch" counterfactual, and generalizing or extrapolating the findings.

In cases where the innovation in question is not embedded in a specific technology with a unique origin, it is challenging to establish that uptake of an observed innovation is indeed attributable to CGIAR and not some other source. For example, Vi Agroforestry (a Swedish NGO) has been promoting agroforestry since the 1980s in Kenya, Rwanda, Tanzania, and Uganda. When adoption data are collected from farmers, there is no immediately obvious trail linking the fodder shrubs, improved fallows, alley cropping, and soil and water conservation practices that farmers carry out in the current time period to these same innovations promoted by Vi a decade earlier. Furthermore, the World Agroforestry Centre's (ICRAF) support to Vi's operations is not immediately observable—for instance, some of the support in the form of trainings and workshops goes back decades. Thus, a process-tracing study was conducted to establish the extent of ICRAF's influence on Vi's programming (LePage Morgan, 2017). This study confirmed the links between Vi Agroforestry's activities, and ICRAF and other CGIAR research. Therefore, the Hughes et al. (2017) study of the impact of a Vi Agroforestry program can be considered an impact study of ICRAF research.

Similarly, a study on dairy hubs—mechanisms to organize farmers into producer organizations and link them to markets—used evidence of ILRI research on the design of informal hubs. International Livestock Research Institute (ILRI) research had showed the importance of upgrading rather than suppressing informal dairy markets. The concept of the "traditional hub," first tested in the East Africa Dairy Development Project (EADD), was the subject of an impact assessment in the More Milk in Tanzania (MoreMilkiT) project (Rao et al., 2016).

Research also influences development programs through their design and implementation. Indeed, 41 percent of the policy claims by CGIAR centers, which Renkow (2018) found to be "well-documented" between 2006 and 2014, were "innovations to the operations and management for government agencies and programs". One of these, the subject of an impact assessment in SIAC (Mills, Nelson, and Achdiawan, 2017), was the Landscape Management for Improved Livelihoods (LAMIL) project. Launched in four reserves of Guinea in 2005 to slow or reverse forest degradation through improved governance and institutions, LAMIL's design was influenced by the Center for International Forestry Research (CIFOR) and ICRAF research that identified the lack of community buy-in as having limited the effectiveness of past projects in the area. As in the ILRI case, CIFOR's involvement as a partner in LAMIL's development supports the argument that the research that influenced the program came from CIFOR.

TANGO International (2017) looked at the influence of the International Food Policy Research Institute's (IFPRI) research on gender and intrahousehold resource allocation on the programming of international NGOs (INGOs). Guided by a theory of change on how research influences program design and implementation, the researchers looked for evidence of influence in documents summarizing the gender policies of the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (OECD-DAC) donors as well as by conducting interviews with donors and NGO staff members. The authors also looked at how the academic literature was evolving in this area to determine whether, in the absence of IFPRI research, other researchers would have arrived at similar findings.

A third way to link impact assessment to development programs is to build impact evaluations into development programs to test the effectiveness of particular technologies, practices, or approaches. Fanzo, Fishman, and Downs (2017) attempted to embed an RCT in the PAPSEN-TIPA[8] project. The project disseminates improved horticulture technologies and equipment in Senegal, including adaptations of the drip irrigation technology co-developed by the International Crops Research Institute for Semi-Arid tropics (ICRISAT), along with complementary vegetable seed inputs and cultivation practices. While the study is likely to provide important information on the links between nutrition-sensitive production and dietary diversity, delays in the PAPSEN-TIPA project resulted in a delay in the RCT implementation, so results are not yet available.

To summarize, despite the potential benefits of collaboration with development programs in scaling and evaluating research, several SIAC impact assessment studies faced methodological and implementation challenges. More rigorous methods, more time allocated toward design and piloting, and/or more experience working with development programs in these types of studies should go some way in helping study teams to produce more and better evidence.

[8] PAPSEN-TIPA: Programme d'Appui au Programme National d'Investissement de l'Agriculture du Sénégal-Technology Adoption for Poverty Alleviation

## 3.4 CONCLUSIONS ON THE SIAC PORTFOLIO

The scale of the studies; the types of commodities, traits, or outcomes assessed; and/or the impact assessment methodologies used in the studies that make up the SIAC portfolio offer new perspectives and add to the existing evidence base. There are a significant number of learning-oriented studies on production systems, which are timely given the growing evidence of low adoption of production systems–related innovations at scale (Stevenson and Vlek, 2018).

Several studies attempted to fill important evidence gaps related to policy and institutional research—such as IFPRI's research on gender and intrahousehold relations and ILRI's work on dairy value chains (Box 2). While methodological or implementation challenges meant that the studies did not result in quantitative estimates of impact, they generated important new knowledge. In general, producing quantitative estimates of the impact of policy-oriented research remains a challenge (Renkow, 2018). While experimental or quasi-experimental methods can be used in some cases, there remain difficulties in attribution as well as in establishing counterfactuals (for example, is the counterfactual the policy that was in place or the new policy but without CGIAR policy-oriented research?). The lessons learned from these studies should inform and spur investment in methodological innovation in these important areas.

In terms of data and methods, the portfolio represents a significant expansion in the approaches that have traditionally been used to assess the impact of agricultural research by SPIA and others. This expansion led to important methodological lessons (refer to Gollin, Probst, and Brower, 2018; Stevenson, Gollin, and Macours, 2018) as well as to new evidence on outcomes and impacts, which is discussed in the sections that follow.

# 4 EVIDENCE OF IMPACTS ON SRF OUTCOMES

Innovations in agriculture are generally expected to have direct effects on yield, input use, and productivity, but they can have broader effects as well. Innovations that are widely adopted and remain in use over many years may have much wider-ranging secondary impacts, including effects on production, wages, the price of the product in question, and the prices of substitute or complementary products. Effects may extend to urban as well as rural consumption and nutrition, to persons living in poverty, and to many aspects of the natural resource environment. Innovations in policy-oriented research may affect such factors as the balance of decision-making between men and women. This section summarizes SIAC evidence on how CGIAR research may have affected such longer-term and broader dimensions of life in target countries.

SPIA recognized that such work would likely depend to some extent on the availability of existing data and that it would pose formidable methodological challenges—not the least of which is the challenge of constructing *ex post* counterfactuals. It was anticipated that different strategies and approaches would be employed, targeting various indicators of impact, with the objective of bringing different types and strength of evidence to bear.

This section summarizes the findings from the 15 studies[9] that generated quantitative evidence related to impact at scale on outcomes defined in the 2016–2030 CGIAR Strategy and Results Framework (SRF) (Table 3) (CGIAR, 2016). In addition to the three System Level Outcomes (SLOs) described earlier, the SRF defined 14 Intermediate Development Outcomes (IDOs) associated with the SLOs and four crosscutting issues (Figures 2 and 3). Because of their importance across multiple IDOs and SLOs, findings related to technology adoption are summarized first, followed by a summary by SLO—poverty (SLO 1), food and nutrition security (SLO 2), and natural resources (SLO 3). When the same study addresses multiple SLOs, the results are presented across multiple sections. Crosscutting issues (Figure 3), for the few studies that measured them, are discussed under the relevant SLO category.

[9] Excluded from this section are the eight learning-oriented studies and the studies that, owing to study design or to problems in implementation, did not produce credible quantitative IDO or SLO estimates over the study period.

**Figure 2.** High-level outcomes in CGIAR SRF

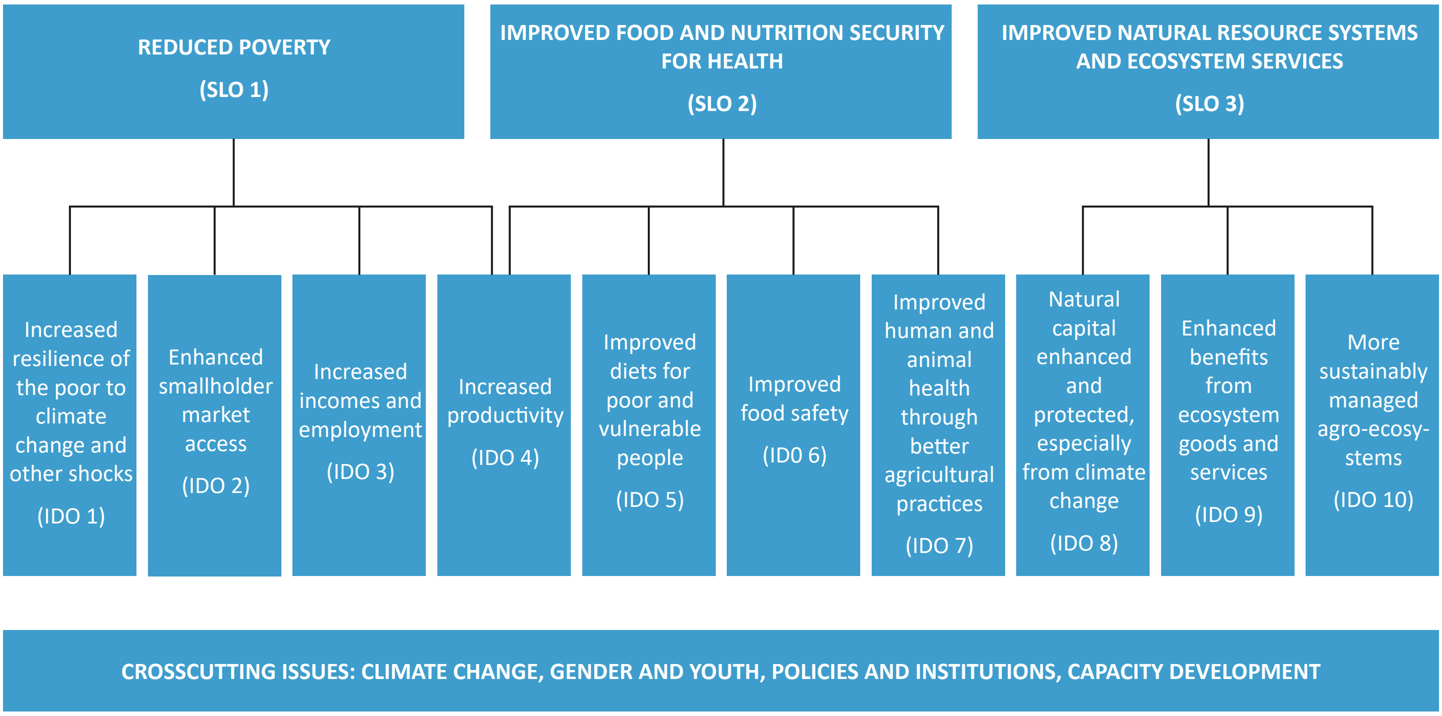
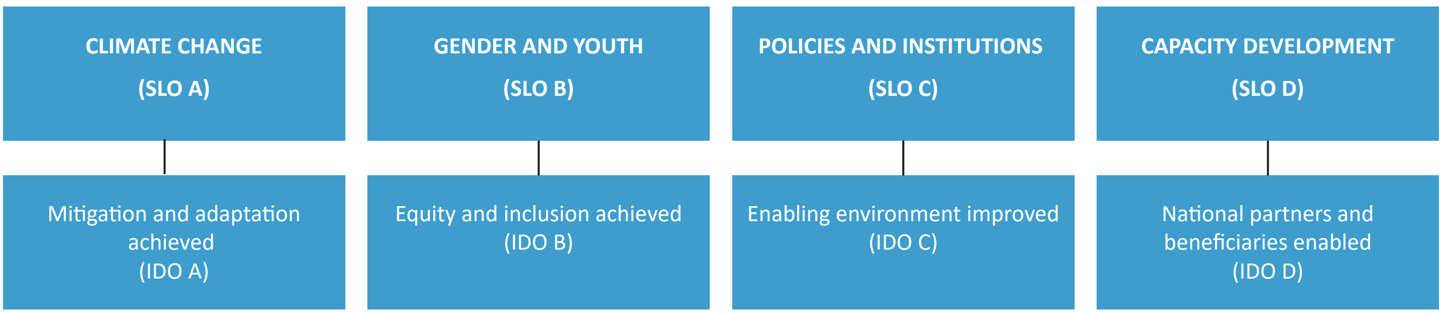



**Figure 3.** Crosscutting SLOs and IDOs




**Source:** Adapted from CGIAR 2016, p. 23.

## 4.1 TECHNOLOGY ADOPTION

Adoption of improved technologies and practices by farmers was central to the impact pathways in 13 of the 15 studies discussed in this section. Seven studies generated new estimates of large-scale adoption of CGIAR innovations (Table 3, column 3). Of these, two studies measured adoption of a specific variety or varieties. Myrick (2016) found that the C88 potato variety, released in 2001 and targeted at farmers producing for the processing market, was grown by 23 percent of potato farmers in Yunnan province in 2015. Labarta et al. (2017a) found that recently released and heavily marketed and promoted high-iron bean varieties have been adopted by 20 percent of farmers in rural Rwanda.

Studies that tracked adoption of multiple CGIAR-related varieties of a particular commodity (released over the years or even across decades) found, as might be expected, higher rates of adoption. Sixty-seven percent of cassava farmers in 16 states of Nigeria, which account for 80 percent of production, grew International Institute of Tropical

Agriculture (IITA)-related varieties (Wossen et al., 2017), and 98 percent of lentil area in western Bangladesh was linked to ICARDA research (ICARDA, 2017). Genetically Improved Farmed Tilapia (GIFT) or GIFT-derived tilapia strains developed by WorldFish and partners were the most commonly produced strains in both Bangladesh and the Philippines, accounting for 53 percent of production in sampled hatcheries in Bangladesh and 40 percent in the Philippines (Benzie and Lind, 2017). While it was not possible to translate adoption by hatcheries to adoption by fish farmers, the authors estimate that the sampled hatcheries account for 25 percent and 22 percent of fry production in Bangladesh and the Philippines, respectively. CIAT-selected *brachiaria* cultivars were planted on 35 percent and 16 percent of pastures in Colombia and Nicaragua, respectively. The uptake of CIAT hybrids,[10] however, was low—less than 0.5 percent (Labarta et al., 2017b). As mentioned, many of these studies identified varieties through DNA fingerprinting, which provides greater accuracy than farmer self-reporting or expert opinion (Box 1).

In a study evaluating an agroforestry development program, Hughes et al. (2017) document a 27 percent increase in adoption of fodder shrubs associated with ICRAF research among dairy farmers in Kenya, against a ten percent increase in the comparison group (dairy farmers in nonprogram areas). The study also constructed an agroforestry adoption index—reflecting both uptake and its intensity—and found that nondairy producers took up agroforestry (both tree and shrub species) more than dairy farmers.

Studies that tracked adoption of multiple varieties and had access to previous adoption surveys not only documented higher numbers of adopters, but also generated a much richer picture of adoption dynamics. For example, the study on lentils in Bangladesh (ICARDA, 2017) showed that more recently released varieties are not as popular as expected, and some of the older varieties are still widely planted. These findings raise important questions for both extension and breeding programs and underscore the importance of having regularly collected data on varieties and varietal turnover in farmers' fields at scales that are sufficiently representative of the population of growers or of significant production acreage. How the adoption data from these studies contributed to estimates of research impacts is discussed in the sections that follow.

A rough estimate based on the five studies that reported adoption numbers is that 3.61 million farm households adopted CGIAR research–related innovations and 4.11 million hectares of crop or grassland area were under CGIAR research–related or CGIAR-selected technologies as of 2015 (Box 3).

## BOX 3. HOW WE ESTIMATED THE CONTRIBUTION TO SRF TARGET #1 ON ADOPTION OF CGIAR-RELATED INNOVATIONS

Based on the five studies that reported adoption estimates, we have attempted to extrapolate the results across the relevant farming populations in the countries of focus for those studies—with the intention of providing evidence on the kind of aggregate numbers targeted in the SRF. The five studies provided new or updated evidence that approximately 3.6 million farm households adopted CGIAR research–related innovations and that 4.1 million hectares of crop or grassland area were under CGIAR research–related or CGIAR-selected technologies for the combinations of crops and countries in question, as of 2015.[11] These estimates are explained as follows, and subject to many caveats as further outlined below.

[10] Breeding of *brachiaria* was accomplished around mid-1990s, but the CIAT hybridization and breeding program was initiated in 1989 by the Forage Project of CIAT.

[11] Except in the case of the *brachiaria* study (Labarta et al., 2017b)—for which data collection occurred in 2017, and the agroforestry study (Hughes et al., 2017)—for which data collection occurred in 2016.

1. **Brachiaria, five countries in Latin America and the Caribbean** (Labarta et al., 2017b): The study is nationally representative. For Colombia and Nicaragua, 1,041 and 480 livestock-producing households were surveyed. For Costa Rica, Honduras, and Peru, expert opinion exercises were held to estimate national and subnational adoption. The study focused on (1) introduced pastures that were selected and disseminated by CIAT research, (2) introduced pastures that were selected and disseminated by non-CIAT research, and (3)genetically improved pastures from the CIAT breeding program, research for which started in 1980s and early 1990s. The study estimated that about 3.93 million hectares of pastures in the five countries is planted with CIAT-selected *brachiarias*.

2. **Cassava, Nigeria** (Wossen et al., 2017): The study is representative of cassava growers in North Central, South East, South West, and South South regions (16 states, which contribute 80 percent of cassava production) of Nigeria. In 2015–2016, 2,500 households were surveyed and DNA fingerprinting was used to identify each unique variety at the plot level. The focus was on all improved cassava varieties released since the 1970s. The study found that 66.7 percent of the sampled households (cassava growers) cultivated improved cassava varieties associated with IITA research, equivalent to about 3.13 million households.

3. **High-iron beans** (Labarta et al., 2017a): The study is representative of bean growers in Rwanda. First, for Season 2015B, a listing exercise collected data from 19,575 households located in 120 villages. Of these 19,575 households, 84 percent grew beans and 20 percent of that 84 percent grew high-iron beans. A subset of 1,397 bean-growing households was resurveyed, and this survey collected data on bean adoption between 2012 and 2015. The focus was on high-iron bean varieties released in 2010 (four varieties) and 2012 (six varieties). The study estimated that about 478,620 households across Rwanda grew high-iron beans at least once in the five years preceding 2015 (inclusive).

4. **Lentils, western Bangladesh** (ICARDA, 2017): The study is representative of lentil-growing households in western Bangladesh, which accounts for 75 percent of Bangladesh's lentil area. The focus was on lentil varieties released after 1995. A survey of 1,000 households located in 52 villages was conducted in the 2015 growing season. DNA fingerprinting of lentil seeds from all lentil fields cultivated by each sample household was used to determine whether a variety was related to research by ICARDA and the Bangladesh Agricultural Research Institute (BARI). The study estimated that about 150,000 hectares of lentil area in rice-lentil systems are under improved varieties released after 1995.

5. **Potato (C88), Yunnan, China** (Myrick, 2016): The study is representative of Yunnan potato producers in the late spring potato season. A survey of 615 households from 41 villages was conducted in July–¬September 2015. The focus was on one specific potato variety released in 2001: Cooperation-88. DNA fingerprinting of samples to confirm the genetic identity of the C88 potato variety was limited to those crops self-identified as C88 by farmers. This study estimated that 33,000 hectares of cropped potato area (55 percent of the area) in winter, and that 23 percent of potato-farming households grew C88 in the 2015 late spring season.

A number of caveats are associated with collating evidence on adoption of improved varieties across countries:

- **The studies are not a snapshot in time per se**. The research that resulted in the improved varieties under study goes back to the 1980s and 1990s, and varieties were then released over time (Table 2, column 4). In only one case (Labarta et al., 2017a) do the varieties under study (released in the second phase of HarvestPlus research) pertain to 2012 and after. Additionally, the data collected in the surveys do not pertain to the same season and year across countries.

- **The quality of data**. Even when the sample of households surveyed was arguably representative at a national or subnational level, data quality may vary between surveys. The studies on lentils, C88 potato, and cassava used DNA fingerprinting for varietal identification whereas the studies on high-iron beans and *brachiaria* relied on farmers' self-reported data or expert opinion, which have been shown to be unreliable in many contexts.

- **Overlap in adoption at the household level**. The five studies from which numbers were drawn were conducted in different countries. It is possible to imagine a case where multiple surveys targeting the same households were conducted within the same country or subnational region, and one would need to ascertain how much of the sample overlapped (i.e., households that adopted multiple CGIAR research–related technologies).

- **Type of data reported**. Studies reported national or subnational adoption by either households or cropped area. Specifically, the studies of lentils, potatoes, and *brachiaria* reported large-scale adoption in terms of hectares, and the other studies in terms of household-level adoption. Interpreting adoption reported in terms of households is tricky since only one plot out of many or a subplot may have been cultivated to that variety.

These caveats illustrate a rationale for carefully planned data collection exercises (at national or subnational levels that occur at sufficiently frequent intervals) and collection of data on multiple CGIAR research–related innovations simultaneously as an alternative to diverse efforts.

## 4.2 PATHWAYS TO POVERTY IMPACTS (SLO 1)

The SRF suggests various pathways from agricultural research to reduced poverty: increased resilience (IDO 1), enhanced market participation and diversification of income sources (IDOs 2 and 3), and increased productivity (IDO 4). Ten of the SIAC studies sought to document impact on poverty-related outcomes through one or more of these pathways (Table 3, column 4). Six studies, all on improved varieties, focused primarily on the productivity pathway. Three studies looked at pathways that combined on-farm productivity with other outcomes, and one looked at a market-based intervention. Of the studies that focused on the productivity pathway to poverty reduction, five collected new data on productivity- and poverty-related outcomes, and one used new techniques to analyze existing data.

### 4.2.1 Improved varieties

In terms of productivity, Wossen et al. (2017) found that adoption of improved cassava varieties (linked to IITA research) was associated with yield gains of 82 percent in Nigeria. The ICARDA study found that improved lentil varieties in western Bangladesh were associated with yield increases of between 356 and 382 kg per hectare (27 percent) compared with pre-1996 landraces (ICARDA, 2017). However, no significant yield gain from varieties released in or after 2006 was found. Myrick (2016), using household harvest and plot area estimates, reported that C88 yielded 24 tons per hectare compared with 19 tons per hectare for a comparable weighted average of all varieties excluding C88—a 26 percent increase.

Wossen et al. (2017) assessed the impact of adoption of improved cassava varieties on a poverty indicator (Foster–Greer–Thorbecke index). Using a combination of econometric and modeling techniques, they found that adoption of improved cassava was associated with a 4.7 percentage-point reduction in poverty—i.e., 1.8 million fewer rural cassava producers in poverty (at US$1.25 per person per day) in 2015–16. Laborde, Martin, and Tokgoz (2017) attempted to assess the contribution to poverty through improved agricultural productivity driven by CGIAR research across 33 countries during 1995-2015. Data and methods-related challenges meant that their results should be interpreted with caution (Box 4). However, addressing these challenges to enable further analyses of this type is an important area for future research.

### BOX 4. CGE MODELS IN *EX POST* IMPACT ASSESSMENT

Given that farmers grow multiple commodities in cropping systems that—in many cases—are integrated with livestock, how would innovations in those other commodities change estimates of the impact of innovation in the commodity under study? How would changes in population, social factors, and nonfarm economic conditions change the impact of these innovations? In theory, a national computable general equilibrium (CGE) approach offers potential for answering such important questions. While CGE models for individual countries face the question of how international trade affects internal prices and quantities, multicountry, multicommodity, dynamic models may help address these issues.

The CGE model in a World Bank–IFPRI study (Laborde, Martin, and Tokgoz, 2017) covered 31 countries, representing about 65 percent of the world's poor. The study uses a multicountry, multisectoral, dynamic CGE model (MIRAGRODEP) to estimate how macroeconomic variables affect production and consumption at the household level. To model the impact of productivity improvements on poverty, the study used both the estimated direct impact of the productivity change on the income of producing households and the estimated impacts of changes in key economic variables, such as commodity prices and wage rates, on the welfare of the household. The resulting distribution of income was used to estimate the change in poverty at the national level. To estimate CGIAR's contribution to the reduction in poverty, the authors used a Delphi approach to obtain subjective estimates (by a panel of observers) of the rates of productivity gains in crops, and the share of those gains attributable to CGIAR research.

The number and diversity of markets, countries, commodities, and innovations make the exercise inherently challenging. Such models incorporate many parameters, any one of which may make a great difference in the implications for any innovations examined. A critical parameter used in the study is total factor productivity (TFP) for crop agriculture as a whole—the size, the length and functional shape of the "lag" associated with innovations' introduction, uptake, and eventual decline. The dramatic difference implied by the results generated by the 30-year polynomial lag and the 50-year gamma function lag used in the World Bank–IFPRI study illustrate the challenges.

Another important parameter for CGE approaches is TFP for individual crops within countries. The authors reported that TFP values used in the modeling exercise were internally derived within the model by combining "historical data and our structural CGE to provide a set of TFP estimates consistent with our theoretical structure, parameters values, and actual evolution." (Laborde, Martin, and Togkoz, 2017, p.16) The procedure provided 297 estimates of crop productivity growth rates per year from 1995 to 2005 for nine crops or crop groups and 33 countries or country groupings. The study also reviewed and presented tables of TFP elasticities with respect to research and development (R&D) expenditures estimated from earlier studies. Given the importance of these TFP parameters, one might ask how sensitive the model results would be to variation in them and in other key parameters of the model.

**Source:** Adapted from Herdt (2018), p.3.

Other studies proxied poverty with intermediate outcomes related to economic benefits from adopting improved varieties. ICARDA (2017) estimated that the higher yields of 1996 or post-1996 lentil varieties, along with reduced costs of labor and pesticides, could have raised farmers' net returns per hectare by 29 percent. Along with the improved consumption traits of varieties released after the late 1990s, this could help explain the almost complete replacement of local landraces by improved lentil varieties in western Bangladesh.

To estimate the economic benefit from C88, Myrick (2016) combined survey data with an expert panel estimate of the time path of adoption between 1996 and 2015, with peak adoption (30 percent of potato land area) estimated to have occurred in 2007. According to the experts, adoption may have declined in recent years because C88 is not as resistant to late blight as previously understood, seed systems have not been responsive to farmer and industry needs, and yields and resistance have declined owing to reuse of seeds. Despite this decline in adoption, the study estimated that C88 could have still generated US$2–3 billion in economic benefits over the 19-year period.

At the macro level, Gollin, Hansen, and Wingender (2018) combined existing data on adoption of CGIAR crop varieties for ten crops and 84 countries over the period 1960–2000 with new data on agro-ecological zones, using econometric and general equilibrium modeling (Box 5). They found that productivity increases—due to improved varieties and related changes—contributed to a 139 percentage-point increase in GDP per capita over the period, with 60 percent of the increase coming from productivity increases and about 40 percent from structural transformation, and in particular from the movement of agricultural workers into occupations that are more productive.

## BOX 5. NEW INSIGHTS ON THE IMPACTS OF THE GREEN REVOLUTION FROM MULTICOUNTRY, MULTICROP STUDIES

Two studies combined existing data sets with new analytical approaches to generate new insights about the impacts of Green Revolution on development outcomes.

- Gollin, Hansen and Wingender (2018) examined adoption data for ten major food crops across 84 countries over 1960–2000—combining global agro-ecological zone (GAEZ) data on agro-climatically attainable crop yields under local climatic conditions, and the spatial/temporal variations in diffusion patterns of high-yielding varieties (HYVs). Newly available data sources such as GAEZ, which indicate geographically determined suitability for HYVs, combined with data on research success made it possible to instrument for the timing and magnitude of productivity shocks received by countries—allowing the authors to model impacts at the national economy (not local labor market) level.

  The estimates suggest that productivity increases from adoption of HYVs and associated changes in agronomic management as well as structural transformation (urbanization, non-agricultural productivity) resulted in a 139 percentage-point contribution to growth between 1960 and 2000. Across the 84 developing countries, average HYV adoption rate was 58 percent, and HYV adoption rates were positively related to life expectancy at birth: a 10 percentage-point increase in HYV adoption was estimated to increase life expectancy by 1.34 percent.

- Fishman et al. (2017) combined the Evenson and Gollin (2003) dataset on adoption of modern varieties (MV) at the country level with datasets on spatially explicit crop harvested area (hence, constructing a proxy for localized MV distribution rates), and infant births and mortality data from the Demographic Health Surveys (DHS) for 18,000 villages across 36 countries and 400 administrative regions. Much like Gollin, Hansen, and Wingender (2018), the ability to examine health outcomes at a global scale, not documented before, was possible because of the availability of geo-referenced data for indicators of interest.

  Results show that the diffusion of modern varieties reduced infant mortality by three to four percentage points (from a baseline of 17 percent), with stronger associations for male infants. The decline in infant mortality translates into about three to five million infant deaths averted per year by the year 2000.

These results emphasize the importance of long-term, large-scale data and of collecting and storing data in ways that enable researchers to link with other data sets from other sectors (such as nutrition and health) and types of data (such as from remote sensing).

### 4.2.2 Production systems

Two studies examined poverty impacts through multiple pathways that include, but are not limited to, increased productivity. In a study on the impacts of conservation agriculture (CA) in Zimbabwe, Michler et al. (2017) focused on increased resilience (IDO 1). The study used existing panel data on yield and inputs for multiple crops (maize, sorghum, millet, groundnut, and cowpea) over four years.[12] The results suggest that CA has little, or even a negative, effect on yields during periods of average rainfall compared with conventional practices. But CA is effective in mitigating the negative impacts of high- and low-rainfall shocks on yields for maize. For sorghum and cowpea, CA appears to mitigate yield loss only during droughts. Since the adoption of CA is also associated with differential input costs (higher fertilizer use but lower seed requirements), the existence of trade-offs between mitigating the effect of shocks (resilience) and yields in normal rainfall years could explain low adoption rates, and has important implications for how this innovation should be targeted and disseminated. Michler et al. (2018) also computed predicted gross revenue to both CA and conventional practice, and the results are fairly consistent with the findings on yields.

Hughes et al. (2017) looked at the impacts of a complex agroforestry intervention informed by ICRAF and CGIAR research—on fodder shrubs, improved fallows, understory multipurpose trees, soil and water conservation, and alley cropping—on a range of household development outcomes in Kenya. Findings suggest that adoption of fodder tree species and associated increases in milk production contributed to improvements in a range of poverty-related indicators such as income and assets. These changes were modest, however, and there was substantial heterogeneity. The study did not find any relationship between project activities and resilience, as measured by an index of coping strategies.

### 4.2.3 Policy

Clements, Alwang, and Achdiawan (2017) assessed the impacts of a CIFOR project on upgrading an Indonesian furniture value chain. Six years after project commencement, the study looked at whether the establishment of a producers' association increased profits more for members than for nonmembers and found that it did not. Similarly, Mills, Nelson, and Achdiawan (2017) found no evidence of higher income flows from the production of maize, groundnuts, tree plantations, or live fencing in Guinean forest reserves where the CIFOR-informed forest co-management project (LAMIL) was implemented compared with reserves where the LAMIL was not implemented.

## 4.3 FOOD SECURITY, HEALTH, AND NUTRITION IMPACTS (SLO 2)

CGIAR's SRF highlights multiple potential pathways from agricultural research to improved food and nutrition security for health (SLO 2). In addition to the indirect pathways through reduced poverty (SLO 1) and improved productivity (IDO 4), there are several direct pathways: improved diet quality (IDO 5) and food safety (IDO 6), and reduction in exposure to disease in the food production (rather than consumption) process (IDO 7). While the larger SIAC portfolio included studies spanning the pathways related to food and nutrition security, the studies whose findings are reported here—that is, the subset of accountability-oriented studies that reported results at scale—focused almost exclusively on the productivity-income pathway, and mainly for staple crops (Table 3).

Two global studies used existing datasets to examine health outcomes of modern variety adoption, a proxy for overall agricultural productivity improvement, associated with the Green Revolution (Box 5). Fishman et al. (2017) found that modern variety adoption between 1960 and 2000 was associated with a significant long-run reduction in infant mortality—at least three to five million infant deaths

[12] Methodologically, using instrumental variables, the study attempted to explicitly control for unobserved heterogeneity among households and selection bias in the choice to adopt CA.

averted per year by the year 2000. Gollin, Hansen, and Wingender (2018) estimate that modern variety adoption contributed to reductions in overall mortality and fertility, resulting in a net negative effect on population growth. These are important new results about the Green Revolution and underscore the fact that it takes a long time and a significant amount of data to accurately measure impacts in the context of long, complex impact pathways. Greater availability of data and greater awareness of the full range of impacts that should be considered in the light of CGIAR SRF and Sustainable Development Goals (SDGs) could shorten this lag.

Two studies that looked at the impacts of improved varieties in the context of maize-legume intercropping found some evidence of improvements in nutrition-related outcomes. In Malawi, the authors found a positive relationship between production diversity (that is, the number of crop species or food groups produced) and dietary diversity; however, the results are modest and context specific (Koppmair, Kassie, and Qaim, 2017). In Ethiopia, maize-legume intercropping and improved maize varieties are associated with greater household food consumption diversity and reduced stunting, which the authors argue is due to the fact that hybrid maize growers are able to purchase diversified diets (Kassie et al., 2016). Wossen et al. (2017) found that adoption of improved cassava varieties in Nigeria was associated with a 24 percent increase in per capita food expenditure and a 17 percent increase in the probability of being food secure but did not affect households' dietary diversity nor consumption of fruits and vegetables.

Mills, Nelson, and Achdiawan (2017) found evidence of higher levels of food security among residents in LAMIL forest reserves compared with control reserves. The authors hypothesize that this arose from greater diversity in the consumption of staples and fruits—consistent with higher adoption of improved maize varieties as well as evidence, albeit weak, of higher adoption of tree plantations and increased diversity of trees grown in LAMIL program area plantations. Hughes et al. (2017) found that a slightly higher percentage of women in the Vi Agroforestry program area had adequate levels of dietary diversity—as measured by the minimum dietary diversity for women (MDD-W; FAO and FHI 360, 2016)—vis-à-vis their counterparts in the comparison areas. While Hughes et al. (2017) report an impact on milk productivity, the authors found no evidence that it contributed directly to the increase in dietary diversity. The study also found no impacts on food security as measured by months of adequate household food provisioning (Bilinsky and Swindale, 2010).

## 4.4 IMPACTS ON NATURAL RESOURCES AND ECOSYSTEM SERVICES (SLO 3)

Agricultural research is theorized to influence environmental outcomes (SLO 3) through several related pathways in CGIAR's SRF: conservation of natural capital (IDO 8), enhanced benefits of ecosystem services (IDO 9), more sustainably managed agroecosystems (IDO 10), and mitigation of and adaptation to climate change (crosscutting IDO A). Three SIAC studies address SLO 3 directly—all in the context of sustainable development programs that are themselves examples of the influence of CGIAR research on policies and institutions (crosscutting IDO C).

Mills, Nelson, and Achdiawan (2017) estimate that a forest co-management project in Guinea (LAMIL), which was informed by CIFOR and ICRAF's policy research, contributed to a reduction in the rate of loss of natural forest, with a value of US$7–14 million in terms of carbon sequestration (aggregated over 2010–2015/16). In their study of impacts of Vi Agroforestry's program in Kenya, Hughes et al. (2017) found that plots in program areas minimally gained in tree cover, but there were no significant changes in total vegetative cover, soil erosion, or soil organic matter.

Flores (2016) looked at the effect of Indonesia's forest moratorium, enacted in 2011 and influenced by research from CIFOR, on forest conservation. While the moratorium bans granting new concessions to oil palm, timber, and logging plantations in all primary forests and in forested and non-forested peatlands, this study focused on peatlands within areas classified as primary forests. The authors used detailed geospatial datasets to compare pre-moratorium (2000–2010) and post-moratorium (2011–2013) peat forest cover for nondesignated areas in comparison with protected areas and concession areas. The study did not find evidence that the moratorium increased forest cover in areas not already classified as protected forests (that is, nondesignated areas, which were expected to gain protection under moratorium) or in concession areas. However, estimates show that after the moratorium was implemented, retained forest cover rose in protected areas relative to rates in concession areas. It is important to note that the study looked at outcomes shortly after the moratorium was implemented, whereas impacts may take longer to emerge, and—in the context of Indonesia—may need to be observed through El Niño cycles.[13]

Overall, two of the three studies found positive relationships between CGIAR research and environmental outcomes. While the impacts were in some cases modest, it is important to note that both the agroforestry study (Hughes et al., 2017) and the forest governance study (Mills, Nelson and Achdiawan, 2017) systematically assessed outcomes across all three SLOs. No other studies in the SIAC portfolio did this, though some hypothesized environmental benefits—for example, the study of CIAT-related *brachiaria* cultivars (Labarta et al., 2017b) or the studies looking at production system changes such as intercropping (Koppmair, Kassie, and Qaim, 2017; Kassie et al., 2016) or relay cropping with lentils (ICARDA, 2017). Gollin, Hansen, and Wingender (2018), in explaining the direct and indirect effects of improved varieties, estimate that the introduction of those varieties is associated with less land use—that is, agricultural land expanded less rapidly in countries where the Green Revolution occurred, consistent with the "Borlaug hypothesis". Few studies in the larger literature examine the synergies and trade-offs between various social and environmental goals or outcomes (Snilstveit et al., 2016), and the ones available suggest that trade-offs may be common (Rasmussen et al., 2018). Whether and how synergies and trade-offs should be addressed in impact assessment is a topic of both academic and policy interest that merits additional research.

[13] The 2011 moratorium on new conversions of primary forest and peat less than 3 meters deep has been renewed several times since. In 2016, a new moratorium was introduced covering all peatlands and requiring companies to restore degraded peatlands. It was reported in 2018 that primary forest loss in Indonesia's protected (peatland) areas dropped 88 percent in 2017 to the lowest level in years (Plumer, 2018).

# 5 IMPACT ASSESSMENT FOR LEARNING: RIGOROUS EVIDENCE ON RESEARCH AND DEVELOPMENT EFFECTIVENESS

The previous section summarized evidence on long-term, large-scale impacts of CGIAR research on development outcomes. While the findings suggest that CGIAR research has made significant contributions, methodological shortcomings often limit the ability to draw clear causal inference, in part because of the long and complex impact pathways between agricultural research and outcomes such as poverty or health. The studies summarized in this section (Table 4) take a different approach, using experimental methods to rigorously assess the effect of research innovations on IDOs (and sometimes SLOs). Because experimental studies often evaluate research outputs in targeted areas before widespread diffusion and purposely manipulate access to and/or information about new technologies, they can be useful for identifying factors that constrain or enhance uptake and impacts and informing the design of scaling-up strategies.

This section reports findings of seven SIAC studies. In three studies, the primary objective was to measure the effect of a CGIAR innovation on a development outcome. Four additional studies focused on understanding factors that enhance or constrain uptake and impact of on-farm soil and water management practices. Because these types of experimental studies are a relatively new approach to better understand impact pathways of CGIAR research, this section includes more methodological detail than section 4.

## 5.1 EVIDENCE ON DEVELOPMENT OUTCOMES

Two SIAC studies sought to rigorously measure the effectiveness of early-maturing rice varieties on development outcomes. Glennerster and Suri (2018) looked at the impacts of NERICA-3 on child undernutrition in Sierra Leone. De Janvry et al. (2017) assessed the impacts of BD-56 on farm profits in Bangladesh. By making the varieties available at the farmers' doorsteps and incentivizing uptake through free or subsidized seeds, the two studies surmount the adoption barrier to be able to observe how farmers and farm households respond to the technologies and make other adjustments in their farm management and livelihood strategies, and to derive implications for outcomes further along the causal chain.

NERICA-3 is an upland, short-duration (90-day) rice variety with high yield potential. Glennerster and Suri (2018) randomly offered NERICA-3 seeds with different subsidy levels and found that take-up was highest (97 percent) in the group whose seeds were fully subsidized; as the subsidy decreased, take-up was lower but still significant (62 percent and 20 percent for the groups who paid half price and full market price, respectively). More importantly, they found that only those farmers who were trained in cultivation practices specific to NERICA-3—in addition to receiving subsidized or free seeds—obtained higher (reported) yields of about 23 percent (45 kg per acre) relative to the control group. No increase in yields was reported by farmers who received seeds but were not exposed to training. Households that were offered training and free seeds[14] also saw positive impacts on a range of child health measures after two seasons—namely, Z-scores for weight-for-height (WHZ)[15] and body mass index for age (BAZ).[16] Without training, although positive impacts on WHZ and BAZ were found, these effects were not significant. That is, accelerated harvest in the absence of higher productivity was an insufficient condition for realizing shorter-term health impacts. Child health measures were stronger for the second season than the first season, possibly because households expanded the area cultivated with NERICA-3. In combination with evidence that rice sales did not increase, the authors hypothesized that impacts are driven by higher productivity rather than early harvest.[17]

In the study on the short-duration rice variety—BD-56—in Bangladesh (de Janvry et al., 2017), the overall take-up rate of mini-kits (5 kg) was 64 percent when seeds were freely distributed. In the experimental control villages, a related variety (BD-51) that is not short duration was distributed,[18] and there was no difference in take-up between BD-56 and BD-51 recipients. BD-56 was found to mature, on average, 25 days earlier than the control variety (BD-51). It also required slightly fewer irrigation days (approximately 0.5 day) in the first, main rice season. The early-maturing trait, however, resulted in lower yields (43 percent lower) in the first, main season after seeds were distributed. Because the shorter cycle offers the possibility of adding an additional cropping season, increasing the number of harvests from two to three (or one to two) could offset reduced yields in normal years. BD-56 plots were indeed 28 percentage points more likely to have a second (*rabi*) crop between the wet and dry seasons and not remain fallow. For the 52 percent of farmers who did cultivate a crop between main agricultural seasons, the annual crop profit was 16 percent higher than for control farmers.[19] Farmers who cultivated BD-56 but kept their land fallow experienced an annual income loss. This, of course, raises the question of why more farmers did not take advantage of the opportunity early maturation presents and plant an additional crop. Since the 5 kg of BD-56 seed distributed would have been sufficient for only about 0.1–0.2 ha (about half the mean plot size), it may not have been feasible to shift that specific part of the plot to a second crop, perhaps for reasons to do with water control or location of the BD-56 area within the plot. Another possibility is that, in the first year of experiment, farmers were not prepared for a *rabi* crop and that more farmers will cultivate a second crop in the following season as they learn more about the variety.

Rejesus et al. (2017) looked at the impact of alternate wetting and drying (AWD) on productivity and income in the Philippines. In contrast to continuous flooding of paddy fields, AWD involves applying water to the field a number of days after ponded water disappears. To help farmers to assess safe water depths, a plastic pipe or bamboo tube is embedded in the paddy field.[20] To assess the impact of AWD,

[14] Consumption and nutrition surveys were conducted only for groups that received the full subsidy (with or without training) because only those arms had sufficiently high take-up rates and were better powered to observe such potential effects.

[15] Wasting—defined in relation to the WHO reference for weight-for-height—is a form of malnutrition that surges during "lean" or "hungry" seasons. Hence, WAZ can change rapidly even in short timeframes as a child gains (or loses) with more (or less) food intake (or an acute illness).

[16] No statistically significant impacts on mid-upper-arm circumference (MUAC) were found.

[17] From this study, one cannot determine whether improved productivity in the absence of accelerated harvest is sufficient to improve child nutrition.

[18] The mini-kits of BD-56 and BD-51 (analogous long duration variety) were sufficient to cover 35 percent of the average wet season rice area.

[19] As it is impossible to know which farmers in the control villages would have added an additional season if given the opportunity, the estimate of the annual profit gain for those who did cannot be interpreted as causal.

[20] The tube is perforated, and soil is removed from inside the tube. The field is irrigated whenever the water level drops below a certain threshold level (as visible inside the tube) pre-determined to be "safe."

farmers in randomly selected irrigation turnout service area groups (TSAGs)[21] were trained on AWD, provided with field water tubes, and encouraged to implement AWD. The authors found no impacts on yields or changes in farmer management practices in the treatment group compared with the control group, and no impact on gross income or water use (Rejesus et al., 2017). The likely reason for the lack of effect is that the conventional practice in the region was similar to AWD. The adoption of AWD by TSAGs/farmers in the TSAGs and pipes on their fields de facto did not alter farmers' ability to manage water and thus offered no substantial benefits.

The NERICA-3 and BD-56 studies provide important evidence on how early-maturing varieties affect farm and household decision making and outcomes. While the studies suggest that impacts on household income and welfare can be positive, they also show that such impacts did not occur automatically after the technologies were introduced, highlighting the challenges in scaling up cost-effectively (for example, the need to invest in extension and training and offer subsidies). More work could be done to understand how impacts vary within households, especially in key areas like labor effort and control over benefits. Such work could help further explain farmers' behavioral responses, including key decisions such as whether to plant an additional crop (as in the case of BD-56).

The AWD study experience underscores the importance of having a good understanding of the control scenario in experimental designs to ensure that it is sufficiently different from the treatment while also allowing isolation of the relevant causal pathway. The study by de Janvry et al. (2017), for example, benefited from both a pilot and trials on similar innovations (flood- and drought-tolerant rice varieties) in the same country as well as in other countries in similar agroecological zones.

## 5.2 EVIDENCE ON DISSEMINATION OF PRODUCTION SYSTEMS INNOVATIONS

Several learning-oriented studies looked at alternative ways to scale up complex production systems management practices: intercropping (Mobarak, 2017), soil fertility analysis–based inputs (Mahajan et al., 2017), and a pilot study on on-farm water conservation (Aker, 2017). These studies took as a starting point that the practices promoted are potentially beneficial but recognized that in light of farmer heterogeneity, what works for whom and how it can best be promoted will vary. This aspect is often not well understood because standard agronomic trials, conducted as part of the technology development process, often do not go beyond plot-level input and output assessment to consider household characteristics and other data. The studies tested different approaches to raising awareness about the practices and to overcoming constraints to adoption. Additionally, the duration of the experiments was established with the aim of understanding whether adoption increased or was sustained over a period of time, reflecting households' perception of benefits (or costs) to adoption.

To adopt a new technology or innovation, farmers must be aware of it and have technical information about it. So, developing an understanding of how best to ease farmers' information constraints or reduce information asymmetry is important (Anderson, 2007). Given the challenges facing extension systems in low- and middle-income countries (Anderson, Feder, and Ganguly, 2006), alternative approaches to extension are needed. These may include incentives for extension workers or changes to extension design (such as the type of information provided or who provides the information), including the leverage of farmers' social networks. In the presence of credit, liquidity, or other market-related constraints, easing just information constraints may not be sufficient to promote adoption—hence, some studies included complementary interventions in the experimental encourage-

[21] Farmers are organized into irrigation associations (IAs) based on the physical layout of irrigation networks. IAs are responsible for operating and maintaining secondary and lateral canals. Within IAs, TSAGs are subgroups that are responsible for turnouts in secondary and lateral canals that then flow to farmers' fields. In the context of the study, farmers in "treated" TSAGs decided when and where to create cracks in the rice bunds to let the water into their paddy fields from the main ditch or lateral canals.

ment design.

Mobarak (2017) experimentally varied whether information about new maize-vegetable intercropping systems was more effectively diffused through agricultural extension workers, lead farmers, or peer farmers in Nepal and compared outcomes with those of control villages.[22] All three types of communicators increased awareness, technical knowledge, and adoption. Adoption increased by 15–30 percentage points across communicator-related treatments compared with the control group in both rounds of midline surveys, although increases in the extension arm were substantially lower in the second round (point estimates). At the time of endline, there was a significant and consistent increase in the adoption of intercropping across the three treatment arms—about 20 percent compared with the control group. There were no statistically significant differences between the treatment arms, though the coefficient on the peer farmer group was lower than on the lead farmer and agricultural extension worker groups. Overall, almost two years after extension, there was some level of sustained adoption—around 20 percent across all treatment arms.[23]

Aker (2017) also found that training on *demi-lunes*, a rainwater-harvesting structure, improved adoption over two seasons in Niger. Training aimed to increase farmer awareness as well as information on demi-lune construction and associated costs and benefits. Since *demi-lunes* are labor intensive and benefit from complementary fertilizer inputs—requiring farmers to incur additional costs—the study also offered cash transfers (conditional and unconditional[24]). Farmers who received unconditional cash transfers and training constructed about 50 percent more *demi-lunes* per hectare compared with farmers who received only training. These farmers also altered their input use—they were 18 percentage points less likely to use inorganic fertilizer and more likely to use organic fertilizer (manure), a complementary demi-lune input, than the training-only group. And although the initial number of *demi-lunes* constructed per hectare was much lower than the number suggested in the training, qualitative follow-up data showed that farmers continued to construct more demi-lunes on the same plots in the following years, perhaps as a result of learning by doing or observation.

Mahajan et al. (2017) studied whether giving farmers customized recommendations on fertilizer dosages, the timing of application, and complementary inputs (such as a precision sowing drill or herbicide) targeted to the soil conditions of their specific plots (one treatment group) or to an average of plots in the village (another treatment group) would improve fertilizer uptake in Mexico. To test for potential liquidity constraints, a random subset of farmers in both groups was also offered subsidized input access (further randomized as flexible or inflexible). Take-up rates exceeded 75 percent for groups that received subsidies but were only seven percent for farmers who did not, and fertilizer use shifted toward recommended practices. Interestingly, results did not differ for the plot-specific versus the village-specific customization for input recommendations, for either uptake or yield. This suggests that, at least in this context, returns to plot-specific soil analysis were limited.

Beyond fertilizer and complementary input use, Mahajan et al. (2017) theorized that being better informed about soil fertility should make farmers less risk averse and allow them to make larger investments to increase productivity. The study found that farmers given localized or individual recommendations were more confident about their own assessment of soil fertility (even if the assessments themselves did not necessarily change). This

[22] While intercropping with the three specific crops promoted (tomato, French beans, ginger) was not as widely known, the study occurred in a context where maize intercropping was widely known and practiced, suggesting low barriers to entry.

[23] Incentives (flat or performance-based) increased adoption by 15–20 percent over the control group, but the "no incentives" arms did better than incentive arms. This anomaly might be due to the delay in the disbursement of incentives after the Nepal earthquake, inducing a negative response from farmers.

[24] Farmers who received conditional cash transfers (that is, conditional on verification of *demi-lunes* per hectare) constructed fewer *demi-lunes* than both training-only and training plus unconditional cash transfers group. The likely reason for this effect is the costs associated with hiring labour for construction of *demi-lunes*—hence, upfront cash transfers had more impact.

interesting result highlights a potentially important mechanism through which access to precise information (through soil analysis) can affect decision making. Consistent with farmers' more accurate assessment of soil fertility, the study finds a lower coefficient of variation (CV) in their yields.[25]

Laajaj and Macours (2017) examined farmers' learning from on-farm experimentation in Kenya. Scientists worked with farmers in villages randomly selected for treatment to implement an experimental trial on their land. Each farmer tested different combinations of seeds (improved soya or maize) and fertilizer packages. By comparing outcomes with those of the control villages, the study investigated how yield gains in the trials translated into farmers' learning. Overall, learning was slow, but farmers selected by the community to participate in the trials were found to learn more and faster (and put in more effort) than randomly selected farmers. These differences, however, decreased over three seasons. While learning increased farmers' willingness to purchase inputs, this was only very incompletely reflected in their actual purchase decisions, suggesting supply-side constraints. These results show the complexity of farmers' learning processes and of translating trial results to real-world settings, and they have implications for the design of extension and diffusion efforts.

The results from this study also show that the way agronomic efficacy trials are set up can significantly influence how researchers develop recommendation domains and characterize the benefits (and costs) associated with a specific technology or innovation (Laajaj et al., 2018). In such trials, researchers play a significant role, and in combination with other aspects of trial design, this external involvement can limit the validity of evidence generated for average farmers under real-life conditions.

These RCTs add to the evidence base on the role access to technical information and farmers' self-learning plays in technology adoption. Providing information on soil conditions (Mahajan et al., 2017) or on technologies and practices (Aker 2017; Mobarak 2017) can increase adoption. At the same time, the type of information provided and who delivers this information can indirectly influence adoption, though the results from SIAC studies are inconclusive. Indeed, Mobarak (2017) finds no significant differences in maize-vegetable intercropping adoption rates between extension workers, lead farmers, and peer farmers in Nepal. Related evidence on the role of social networks and the effectiveness of targeting certain farmers within communities as communicators is rapidly emerging (Magruder, 2018) and may suggest ways to help shorten the time between release and widespread diffusion. Mahajan et al. (2017) further show that it may not always be necessary to target input recommendations and farm management advice at the individual farmer or farm level, making such approaches more scalable.

[25] Since the extension service included recommendations for use of machinery, hybrid seeds, herbicide, and fertilizer at sowing (with or without grants), it is difficult to isolate the effects of extension from a response to increased confidence in soil fertility assessment.

# 6 DISCUSSION

This report has characterized and described the findings from 25 impact assessments funded under SIAC. The studies span a wide range of innovations and scientific knowledge within three main areas of CGIAR research—crop improvement, production systems management, and policy. Studies used a range of methods, from RCTs to qualitative approaches, depending on the specific research question being addressed. Some studies used mixed methods. The majority of the studies looked at impacts of improved varieties and at outcomes related to poverty and food security. That said, the range of crops covered went well beyond the major cereals, and the results add to the evidence base on nutrition and health outcomes and, to a more limited extent, environmental outcomes of agricultural research. This high-quality evidence of adoption and impact at scale, rather than for a small number of villages or pilot sites, is a significant addition to the evidence base.

## 6.1 IMPACTS OF CGIAR-RELATED RESEARCH ON IDOS/SLOS AND ASSOCIATED ASPIRATIONAL TARGETS: KEY RESULTS

**Technology adoption**. Five studies generated new evidence on technology adoption at scale, broadening the base of commodities covered. Use of DNA fingerprinting was demonstrated to be feasible at scale and, in nearly all cases, necessary to accurately estimate adoption and impact. Such analyses also illustrated how DNA fingerprinting improves the quality of evidence on correlates of adoption. One study documented the adoption of agroforestry and other natural resource management practices in the context of a long-term development program. A rough estimate is that SIAC results provide new or updated evidence that 3.61 million farm households were using CGIAR research–related innovations and that 4.11 million hectares of crop or grassland area were under CGIAR research–related or CGIAR-selected technologies in 2015/2016 (see Box 3).

**Poverty (SLO 1)**. Only one study tried to measure the impact of improved technology on poverty, but six nationally or subnationally representative studies report positive impacts of improved varieties and production system interventions on intermediate outcomes such as income, consumption, or assets. At the global level, productivity increases resulting from the adoption of modern varieties associated with CGIAR contributed to a 139 percentage-point increase in overall GDP per capita between 1960 and 2000 in a sample of 84 countries, in part by allowing for structural transformation.

**Food security and nutrition for health (SLO 2)**. Two multi-country studies provide important new evidence on the nutrition, health, and demographic impacts of the Green Revolution.

Farmers' adoption of modern varieties between 1960 and 2000 was associated with a significant long-run reduction in infant mortality (3–5 million infant deaths averted per year by the year 2000) and contributed to reductions in both overall mortality and fertility, with a net negative effect on population growth.

Studies of more recent impacts suggest that changing crop varieties and production practices has some potential to contribute to improved food security and nutrition status, particularly within producer households. The evidence is not conclusive, however, and the evidence base needs to be broadened to include the nutrition and health impacts of market and governance interventions (Bulte, 2018).

**Natural resources and ecosystem services (SLO 3)**. Two studies that looked at development programs informed by CGIAR research document positive environmental outcomes—in one study, a reduction in the rate of loss of natural forest and in another an increase in tree cover. Many other studies with potential natural resource or environmental implications did not measure such outcomes, even when the studies hypothesized that they existed. Therefore, SIAC studies likely significantly underestimate the impacts—positive and negative—of CGIAR research on SLO 3. Quantifying the impacts of CGIAR research on environmental outcomes therefore remains an important area for further inquiry.

**Understanding impact pathways**. Results from RCTs provided important evidence on how early-maturing rice varieties or complex production system management practices affect farm and household decision making and outcomes. They demonstrate that gains in household welfare did not occur automatically after the technologies were introduced, highlighting the challenges of scaling up cost-effectively. In many of the learning-oriented studies, explicit or implicit subsidies or direct access to innovations facilitated additional adoption, which suggests that important constraints (including lack of reliable input supply chains) remain and need to be addressed.

The results also point to the fact that increasing awareness of technologies or enabling access to inputs ought to be based on a good understanding of the returns to such technologies for farmers in real-world settings. This is even more important in cases where CGIAR research improves upon a technology or practice that farmers already use, as in the case of AWD in Philippines or inter-cropping in Nepal. Hence, during efficacy trials and experimental on-farm trials, it is crucial to increase efforts to measure productivity and other outcomes in ways that reflect real-world conditions and the possible behavioral responses of targeted farmers.

## 6.2 METHODOLOGICAL LESSONS FROM THE SIAC PORTFOLIO

**The portfolio illustrates improved rigor and innovative use of tools for quasi-experimental methods and modeling approaches, but more could be done**. Quasi-experimental and modeling methods were the most commonly used approach in SIAC studies. This is not surprising considering the framing of research questions, which focused on quantifying SLO- or IDO-level impacts at scale. Despite the widely-acknowledged identification issues, these approaches are often necessary to estimate long-term impacts that go beyond plot- or farm-level productivity, and they can speak to the kinds of impacts many stakeholders are interested in. Improving the quality of data and methods and increasing the number of studies—so that results of studies of similar innovations conducted in different contexts using different methods can be compared and synthesized—will help strengthen the evidence base (Herdt, 2018).

Some studies benefited from local knowledge (through long-term partnerships) or older process evaluations that documented where and how diffusion had occurred (through public sector or non-profit-led development projects) and how CGIAR research had influenced these development projects. However, such information is not readily and widely available for other contexts. A systematic effort to carefully document both research influence and diffusion efforts, where feasible, could go a long way in informing better impact assessment design.

**Environmental outcomes were hypothesized but rarely documented**. With the exception of studies on innovations related to natural resource management or policy, studies of crop improvement or production system innovations rarely measured environmental outcomes. This was the case even when such outcomes were part of the research theory of change.

**Little attention was paid to synergies and trade-offs among outcomes**. While some SIAC studies, in particular natural resource management studies, were designed to examine the trade-offs that households face, most did not do so systematically. A few studies found evidence of potential trade-offs, especially related to risk management (Michler et al., 2017; Glennerster and Suri, 2018; de Janvry, 2017), which has important implications for appropriate dissemination strategies. Impact pathways should be screened more systematically at an early stage to identify potential trade-offs and address them during research or impact evaluation.

Adoption of an innovation is prima facie evidence that a decision maker considers the innovation beneficial, and continued use of the innovation can be interpreted as evidence that the decision maker believes it is better than the available alternatives. However, that belief is—at best—weak evidence of the innovation's impact on well-being. Using adoption data on particular innovations to make inferences about how innovations impact people's lives presents limitations that are partly related to the synergies and trade-offs that occur on the farm and beyond it. Future studies should therefore incorporate careful consideration of impact pathways, and the possible private and public benefits or costs associated with adoption of particular innovations, including—where appropriate—impacts on wider household activities. Such an approach could also improve the policy relevance of studies, since policymakers and practitioners are often interested in the costs and benefits of innovations in various biophysical and socioeconomic contexts.

**Very little attention was devoted to gender and wider socioeconomic heterogeneity**. With a few exceptions, the SIAC studies did not consider gender or other relevant socioeconomic dimensions. Even disaggregation of outcomes by sex or other relevant socioeconomic dimensions (such as age, caste or ethnic group, farm size or type, farmers' education levels, wealth status) was done in only a few studies. More attention should be paid to ensuring that heterogeneity is well addressed in future CGIAR impact assessments. In cases where specific interventions are discussed, it is unclear whether the findings can be generalized (even in a limited manner) to different types of households or farmers within households.

**More attention should be paid to external validity**. The idiosyncratic nature of the selection of the learning-oriented studies limited the ability to extrapolate findings to other contexts or innovations. Increased coordination between studies at the design stage can help enhance the potential for synthesis and drawing of generalizable lessons. Substantive preliminary work, including careful piloting, is important to assure that such studies lead to broadly relevant lessons.

## 6.3 LESSONS ON DATA AND MEASUREMENT

It is evident that the availability of adoption data at large, representative scales over time that can be linked to other datasets, including but not limited to socioeconomic data could help facilitate impact studies that capture the full range of system-level impacts of CGIAR research.

When the objective is to characterize adoption of crop varieties and their impacts, valid data on both the genetics and the production inputs are needed. The SIAC results on DNA fingerprinting, which are being summarized in a scientific publication and a practitioners' guide, show that in many cases DNA fingerprinting is essential to obtaining accurate adoption estimates. Even where it may be feasible to get accurate aggregate modern variety adoption estimates from, for example, self-reported data, observed type I and type II errors in self-reported varietal data illustrate why fingerprinted data are better for analyzing correlates of adoption. However, practical issues like sampling, preserving tissue samples from deterioration, and generating reference libraries all require careful attention (particularly in developing-country field conditions) to ensure that DNA fingerprinting is carried out correctly.

The use of remotely sensed data and geospatial tools to document adoption and impact shows much promise. It can reduce the cost of data collection, improve the credibility of studies where a baseline is absent and must be created *ex post*, and of course allow association of datasets with other data that are not available at required spatial scales (for example, when rain gauges are not available at the village level) or are costly to collect (for example, data on soil fertility). An implicit assumption here is that such geospatial data and associated models have been ground-truthed and validated (or can be validated as a part of the study).

Finally, for all innovations, there is a need for better measures of productivity and broader measures of potential benefits and costs beyond yields. While many studies documented adoption, less attention was paid to careful measurement of land and labor productivity, income, or profitability. This information is essential to understand how farm households benefit (or do not benefit) from particular innovations, which has implications for predicting adoption and for convincingly linking it to poverty reduction or other outcomes.

# REFERENCES


Anderson, J. R. (2007). Agricultural advisory services. Background paper for *World Development Report 2008*. Washington, DC: World Bank.

Anderson, J. R., Feder, G., and Ganguly, S. (2006). *The rise and fall of training and visit extension: An Asian mini-drama with an African epilogue*. Policy Research Working Paper 3928. Washington, DC: World Bank.

Baylis, K., Honey-Rosés, J., Börner, J., Corbera, E., Ezzine-de-Blas, D., Ferraro, P. J., Lapeyre, R., Persson, U.M., Pfaff, A., and Wunder, S. (2016). Mainstreaming impact evaluation in nature conservation. *Conservation Letters, 9*(1), 58–64. doi:https://doi.org/10.1111/conl.12180.

Beegle, K., Carletto, C., and Himelein, K. (2012). Reliability of recall in agricultural data. *Journal of Development Economics, 98*(1), 34–41. doi:https://doi.org/10.1016/j.jdeveco.2011.09.005.

Bilinsky, P., and Swindale, A. (2010). *Months of adequate household food provisioning (MAHFP) for measurement of household food access: Indicator guide (v.4)*. Washington, DC: FHI 360/Food and Nutrition Technical Assistance II Project (FANTA).

Bulte, E. (2018). *Assessing the Impacts of International Agricultural Research on Nutrition and Health in Africa and Asia*. SPIA Technical Note No. 1. Rome: Independent Science and Partnership Council. https://ispc.cgiar.org/publications/assessing-impacts-international-agricultural-research-nutrition-and-health-africa-and.

Cameron, D. B., Mishra, A., and Brown, A. N. (2016). The growth of impact evaluation for international development: How much have we learned? *Journal of Development Effectiveness, 8*(1), 1–21. doi:https://doi.org/10.1080/19439342.2015.1034156.

Carletto, C., Gourlay, S., Murray, S., and Zezza, A. (2016). *Cheaper, faster, and more than good enough: Is GPS the new gold standard in land area measurement?* Policy Research Working Paper 7759. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/24840.

CGIAR. (2016). *Strategy and results framework 2016–2030: Redefining how CGIAR does business until 2030*. Montpellier, France.

CGIAR Science Council. (2009). *Stripe review of social sciences in the CGIAR*. Rome: Science Council Secretariat.
https://ispc.cgiar.org/sites/default/files/ISPC_StripeReview_SocialScience.pdf.

Duflo, E., and Kremer, M. (2003). *Use of randomization in the evaluation of development effectiveness*. Paper presented at the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness, July 15–16, Massachusetts Institute of Technology, Cambridge, MA, USA. http://www.gsdrc.org/document-library/use-of-randomization-in-the-evaluation-of-development-effectiveness/.

Elven, S., and Krishnan, L. (2018). *Estimating historical CGIAR research investments*. SPIA Technical

Note No. 5. Rome: Independent Science and Partnership Council.

Evenson, R.E. and Gollin, D. (2003). Assessing the Impact of the Green Revolution, 1960 to 2000. *Science*, 300, 758–762. doi:https://dx.doi.org/10.1126/ science.1078710

Evenson, R. E., and Rosegrant, M. (2003). The economic consequences of crop genetic improvement programmes. In R. E. Evenson and D. Gollin (eds.), *Crop variety improvement and its effect on productivity*, pp. 473–497. Wallingford, UK: CABI.

FAO and FHI 360. (2016). *Minimum Dietary Diversity for Women: A Guide for Measurement*. Rome: Food and Agricultural Organization of the United Nations (FAO).

Garbero, A., Marion, P., and Brailovskaya, V. (2016). *Meta-analysis: The impact of agricultural research on poverty*. Statistics and Studies for Development Division (SSD) Working Paper. Rome: International Fund for Agricultural Development (IFAD).

Gollin, D., Probst, L. T., and Brower, E. (2018). *Assessing poverty impacts of agricultural research: Methods and challenges for the CGIAR*. Rome: Independent Science and Partnership Council.

Gryseels, G., de Wit, C. T., McCalla, A., Monyo, J., Kassam, A., Craswell, E., and Collinson, M. (1992). Setting agricultural research priorities for the CGIAR. *Agricultural Systems*, 40(1), 59–103. doi:https://doi.org/10.1016/0308-521X(92)90016-H.

Herdt, R. W. (2018). *Documenting the impact of widely-adopted CGIAR research innovations*. SPIA Technical Note No. 3. Rome: Independent Science and Partnership Council. https://ispc.cgiar.org/publications/documenting-impact-widely-adopted-cgiar-research-innovations.

ISPC (Independent Science and Partnership Council). (2017). Strengthening impact assessment in the CGIAR (SIAC), 2013–2017. https://ispc.cgiar.org/workstreams/impact-assessment/projects/strengthening-impact-assessment-cgiar-siac-2013-2017.

Jutzi, S. C., and Rich, K. M. (2016). *An evaluation of CGIAR centers' impact assessment work on livestock-related research (1990–2014)*. Rome: Standing Panel on Impact Assessment and Independent Science and Partnership Council. https://ispc.cgiar.org/sites/default/files/pdf/SIAC-3.3_livestock-IAs_review.pdf.

Kelley, T., Ryan, J., and Gregersen, H. (2008). Enhancing ex post impact assessment of agricultural research: The CGIAR experience. *Research Evaluation, 17*(3), 201–212.

Kerr, J. M., and Kolavalli, S. (1999). *Impact of agricultural research on poverty alleviation: Conceptual framework with illustrations from the literature*. EPTD Discussion Paper No 56. Washington, DC: International Food Policy Research Institute.

Kilic, T., and Sohnesen, T. P. (2015). *Same question but different answer: Experimental evidence on questionnaire design's impact on poverty measured by proxies*. Policy Research Working Paper 7182. Washington, DC: World Bank. https://ssrn.com/abstract=2579881.

LePage Morgan, C. (2017). From the lab to the farm: Using process-tracing methods to evaluate impact pathways in agroforestry research. Unpublished technical report submitted to the World Agroforestry Center.

Magruder, J. (2018). An assessment of experimental evidence on agricultural technology adoption in developing countries. *Annual Review of Resource Economics 10*, 299–316.

Meinzen-Dick, R., Adato, M., Haddad, L., and Hazell, P. (2004). *Science and poverty: An interdisciplinary assessment of the impact of agricultural research*. Food Policy Report. Washington, DC: International Food Policy Research Institute.

Merrey, D. J. (2015). *An evaluation of CGIAR centers' impact assessment work on irrigation and water management research*. Rome: Standing Panel on Impact Assessment and Independent Science and Partnership Council.

Pingali, P., and Kelley, T. (2007). The role of international agricultural research in contributing to global food security and poverty alleviation: The case of the CGIAR. In R. Evenson and P. Pingali (eds.), *Handbook of agricultural economics*, Vol. 3, pp. 2381–2418. Amsterdam: Elsevier.

Plumer, B. (2018). Tropical forests suffered near-record tree losses in 2017. *New York Times*, June 27. https://www.nytimes.com/2018/06/27/climate/tropical-trees-deforestation.html.

Pray, C. E., Masters, W. A., and Ayoub, S. (2017). *Impacts of agricultural research on poverty, malnutrition, and resilience*. Report prepared for the U.S. Agency for International Development (USAID) Bureau for Food Security, Office of Agriculture, Research and Policy. https://sites.tufts.edu/willmasters/files/2017/04/AgRschImpactsOnPovertyNutritionResilience_7April2017.pdf.

Raitzer, D. A., and Kelley, T. G. (2008). Assessing the contribution of impact assessment to donor decisions for international agricultural research. *Research Evaluation*, 17(3), 187–199. doi:https://doi.org/10.3152/095820208X331702.

Rao, X., Hurley, T. M., and Pardey, P. G. (2016). *Are agricultural R&D returns declining and development dependent?* Paper presented at the Agricultural and Applied Economics Association (AAEA), 2016 Annual Meeting, July 31–August 2, Boston, MA, USA. http://ageconsearch.umn.edu/record/235962.

Rasmussen, L. V., Coolsaet, B., Martin, A., Mertz, O., Pascual, U., Corbera, E., . . . Ryan, C. M. (2018). Social-ecological outcomes of agricultural intensification. *Nature Sustainability*, 1(6), 275–282. doi:https://doi.org/10.1038/s41893-018-0070-8.

Renkow, M. (2018). *A reflection on impact and influence of CGIAR policy-oriented research*. Rome, Italy: Standing Panel on Impact Assessment and Independent Science and Partnership Council.

Renkow, M., and Byerlee, D. (2010). The impacts of CGIAR research: A review of recent evidence. *Food Policy, 35*(5), 391–402.

Savedoff, W. D., Levine, R., and Birdsall, N. (2006). *When will we ever learn?: Improving lives through impact evaluation*. Washington, DC: Center for Global Development.

Snilstveit, B., Stevenson, J., Villar, P. F., Eyers, J., Harvey, C., Panfil, S., Puri, J., and McKinnnon, M. C. (2016). *Land-use change and forestry programmes: Evidence on the effects on greenhouse gas emissions and food security*. Evidence Gap Map Report 3. London: International Initiative for Impact Evaluation (3ie).

Stevenson, J., Gollin, D., and Macours, K. (2018). *The rigor revolution in impact assessment: Implications for the CGIAR*. Rome: Independent Science and Partnership Council.

Stevenson, J., Villoria, N., Byerlee, D., Kelley, T., and Maredia, M. (2013). Green Revolution research saved an estimated 18 to 27 million hectares from being brought into agricultural production. *Proceedings of the National Academy of Sciences, 110*(21), 8363–8368. doi:https://doi.org/10.1073/pnas.1208065110.

Stevenson, J., and Vlek, P. (2018). *Assessing the adoption and diffusion of natural resource management practices: Synthesis of a new set of empirical studies*. Rome: Independent Science and Partnership Council.

Waibel, H., and Zilberman, D. (eds.). (2007). I*nternational research on natural resource management: Advances in impact assessment*. Rome: CABI International and Food and Agriculture Organization of the United Nations.

Webb, P. W. (2013). *Impact pathways from agricultural research to improved nutrition and health: Literature analysis and research priorities*. Rome and Geneva: Food and Agriculture Organization of the United Nations and World Health Organization. http://www.fao.org/fileadmin/user_upload/agn/pdf/Webb_FAO_paper__Webb_June_26_2013_.pdf.

# SIAC-FUNDED IMPACT ASSESSMENTS

Aker, J. (2017). *Are rainwater harvesting techniques profitable for small-scale farmers? A study on the impacts of adopting rainwater harvesting techniques in Niger*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Benzie, J., and Lind, C. (2017). *Adoption of Genetically Improved Farmed Tilapia (GIFT) twenty years after release to industry: A GIFT that keeps on giving?* Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

CIP (International Potato Center). (2017). *Adoption and diffusion of C88 potato variety in China: Spatial variability of productivity gains and cost savings and value chain development*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC). Lima, Peru: CIP.

Clements, C., Alwang, J., and Achdiawan, R. (2017). *Upgrading the Jepara furniture value chain*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

de Janvry, A., Emerick, K., Sadoulet, E., and Dar, M. (2017). *Drought resistance and water saving in rice production in Bangladesh*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Fanzo, J., Fishman, R., and Downs, S. (2017). *Agriculture, women time use, income and nutrition linkages: Leveraging a randomized control trial in Senegal*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Fishman, R., Dar, A., Barnwal, P., Mueller, N. D., von der Goltz, J., and McCord, G. C. (2017). *Crop variety improvement and infant mortality in the developing world*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Flores, N. L. (2016). *Impact assessment of natural resource management policy research: A case study of the contribution of the Sustainable Wetlands Adaptation and Mitigation Project to the effectiveness of the Indonesian forest moratorium*. Master's thesis, Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA, USA. http://hdl.handle.net/10919/71894.

Glennerster, R., and Suri, T. (2018). *Agricultural technology and nutrition: The impacts of NERICA rice in Sierra Leone*. Working Paper. http://mitmgmtfaculty.mit.edu/tsuri/working-papers/.

Gollin, D., Hansen, C. W., and Wingender, A. M. (2018). *Two blades of grass: The impact of the Green Revolution*. NBER Working Paper No. 24744. Cambridge, MA: National Bureau of Economic Research. DOI: https://doi.org/10.3386/w24744.

Hughes, K., Morgan, S., Baylis, K., Oduol, J., Smith-Dumont, E., Vagen, T.-G., Mutemi, M., LePage, C., and Kegode, H. (2017). *Assessing the downstream socioeconomic and land health impacts of agroforestry in Kenya*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC). A revised version of the report is available at http://www.worldagroforestry.org/downloads/Publications/PDFS/WP18033.pdf.

ICARDA (International Center for Agricultural Research in the Dry Areas). (2017). *Adoption of improved lentil varieties in Bangladesh: Comparison between expert estimates, nationally representative farm household survey and DNA fingerprinting*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC). Beirut: ICARDA.

Kassie, M., Marenya, P., Teklewold, H., Erenstein, O., Matin, Q., and Dil, R. (2017). *Adoption of cropping system diversifications and modern seeds affecting nutrition security in Ethiopia*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC) and paper presented at the 30th International Association of Agricultural Economists (IAAE), July 28–August 2, 2018, Vancouver, British Columbia. https://ideas.repec.org/p/ags/iaae18/277170.html.

Koppmair, S., Kassie, M., and Qaim, M. (2017). Farm production, market access, and dietary diversity in Malawi. *Public Health Nutrition, 20*(2), 325–335. doi:https://doi.org/10.1017/S1368980016002135.

Laajaj, R., and Macours, K. (2017). *Learning-by-doing and learning-from-others: Evidence from agronomical trials in Kenya*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Laajaj, R., Macours, K., Masso, C., Thuita, M., and Vanlauwe, B. (2018). *Yield gap or field gap? Reconciling yield gains in agronomic trials and African smallholder conditions*. Unpublished paper, Universidad de Los Andes, Bogota.

Labarta, R. A., Birachi, E., Katungi, E., Vaiknoras, K., Larochelle, C., Alwang, J., Asare-Marfo, D., Herrington, C., and Birol, E. (2017a). *Assessing impacts of the adoption of high iron bean varieties on bean and iron intake of rural vulnerable populations in Rwanda*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Labarta, R., Martinez, J. M., Yaccelga, A., Reyes, B., Gomez, L., Maredia, M., DeYoung, D., Carriazo F., and Toro, M. (2017b). *Assessing the adoption and economic and environmental impacts of Brachiaria grass forage cultivars in Latin America, focusing on the experience in Colombia*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Laborde, D., Martin, W. and Tokgoz, S. (2017). *Poverty impacts of widely-adopted CGIAR innovations*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Mahajan, A., Seira, E., Gine, X., and Corral, C. (2017). *A multiple intervention approach to increasing technology adoption with a view towards scaling-up: Evidence from Mexico*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Michler, J.D., Baylis, K., Arends-Kuenning, M., and Mazvimavi, K. (2017). *Conservation agriculture and climate resilience*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Michler, J.D., Baylis, K., Arends-Kuenning, M., and Mazvimavi, K. (2018). Conservation agriculture and climate resilience. *Journal of Environmental Economics and Management, 93*, 148-169. https://doi.

org/10.1016/j.jeem.2018.11.008.

Mills, B., Nelson, C., and Achdiawan, R., (2017). *Into the forest with or without a trace? A multi-level impact analysis of forest co-management in Guinea*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Mobarak, M. (2017). *Using social networks to promote new agricultural technologies*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Myrick, S. (2016). *An economic impact assessment of Cooperation-88 potato variety in the Yunnan Province of China*. M.S. thesis, Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA, USA.

Rao, J., Baltenweck, I., Girard, A., Dominguez-Salas, P., Galie, A., and Poole, J. (2016). *Looking beyond income: Impact of dairy hubs on human nutrition in Tanzania*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

Rejesus, R. M., Yorobe Jr., J. M., Lampayan, R. M., Sibayan, E. B., Ole Sander, B., Mendoza, M. L., B. Flor, R. J., B. Malabayabas, A. J., Singleton, G. R., Mohanty, S., and San Valentin, M. R. (2017). *Adoption and impact of alternate wetting and drying (AWD) water management for irrigated rice production in the Philippines*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

TANGO. (2017). *Assessing the impact of IFPRI's intra-household research program, 1995–2000, on the field practice of INGOs*. Unpublished report submitted to the Standing Panel on Impact Assessment of the Independent Science and Partnership Council (ISPC).

Wossen, T., Abdoulaye, T., Alene, A., Feleke, S., Rabbi, I., and Manyong, V. (2017). *Assessing the impact of improved cassava varieties on food security and poverty reduction in Nigeria*. Unpublished report submitted to the Standing Panel on Impact Assessment (SPIA) of the Independent Science and Partnership Council (ISPC).

**Table 2.** SIAC-funded studies on impacts of CGIAR research

| Innovation | SIAC call window[a] | Research area | Specific research assessed | Methods | IDOs and SLOs targeted[b] |
|---|---|---|---|---|---|
| 1. Better understanding of gender relations (*TANGO 2017*) | LTLS | Policy | IFPRI's gender and intrahousehold research program ran from 1996 to 2005. | Qualitative (document review and key informant interviews) | Gender equity (CCT) |
| 2. Improved cassava varieties in Nigeria (*Wossen et al., 2017*) | LTLS | Germplasm enhancement and breeding | IITA started working on cassava varietal improvement in the early 1970s, focusing on disease resistance and yield. The first variety was released in 1977. | Quasi-experimental (cross-sectional survey) | Adoption<br>Food security and nutrition (SLO 2)<br>Productivity (SLO 1) |
| 3. Improved lentil varieties in Bangladesh (*International Center for Crop Research in Dryland Areas [ICARDA], 2017*) | LTLS | Germplasm enhancement and breeding | Research to develop new lentil varieties began in the mid-1980s in response to a production crisis (disease susceptibility). The first variety (BARI-1) was released in 1991. Study focuses on varieties released after 2006, including BARI-5 and BARI-6. | Quasi-experimental (cross-sectional survey) | Productivity (SLO 1) |
| 4. Improved potato variety C88 in China (*International Potato Center [CIP], 2017 and and Myrick, 2016*) | LTLS | Germplasm enhancement and breeding | Starting in the mid-1980s in response to late blight, CIP and Yunnan Normal University (YNNU) collaborated to develop a late blight–resistant potato variety. Research ran from 1986 to 1996, and Cooperation 88 (C88) was released in 2001. | Macro-level economic impact assessment (economic surplus) | Productivity, economic gains (SLO 1) |
| 5. Improved tilapia (fish strains) in Philippines and Bangladesh (*Benzie and Lind, 2017*) | LTLS | Germplasm enhancement and breeding | GIFT was introduced in 1994, and research occurred even earlier (1988). Study examined adoption 20 years after release. | DNA fingerprinting (cross-sectional survey) | Adoption |
| 6. Multi-crop, multi-country (*Fishman et al., 2017*) | LTLS | Germplasm enhancement and breeding | Modern varieties from CGIAR, 1959–2001 (36 countries) | Macro-level economic impact assessment | Health (SLO 2): infant mortality |
| 7. Multi-crop, multi-country (*Laborde, Martin, and Tokgoz, 2017*) | LTLS | Germplasm enhancement and breeding | Modern varieties from CGIAR, 1994–2015 data | Macro-level economic impact assessment (computable general equilibrium model) | Income, poverty (SLO 1) |
| 8. Improved value chain: Dairy hubs to improve productivity and market access (*Rao et al., 2016*) | Nutrition | Policy | The dairy hub concept grew out of ILRI's smallholder dairy project (1997–2005) and subsequent dairy value chain work. The second phase of the East Africa Dairy Development project started in 2014; new hubs (for this study) were set up as a part of this. | Quasi-experimental (propensity score matching) | Nutrition (SLO 2) |
| 9. Improved high-iron beans in Rwanda (*Labarta et al., 2017a*) | Nutrition | Germplasm enhancement and breeding | Second wave of five bean varieties with higher iron content released in 2012. The research is part of HarvestPlus, which started in 2003. | Quasi-experimental (propensity score matching) | Nutrition (SLO 2): iron intake<br>Income (SLO 1) |
| 10. Improved (short-duration) rice (*Glennerster and Suri, 2017*) | Nutrition | Germplasm enhancement and breeding | New Rice for Africa (NERICA) research started in the late 1990s, and by 2000, varieties had been released in AfricaRice member countries. Intense dissemination began in 2005. This study focuses on NERICA3, an upland variety with early maturity. | Experimental | Productivity (SLO 1)<br>Food security, nutrition (SLO 2): wasting, body mass index, mid-upper-arm circumference |

| Innovation | SIAC call window[a] | Research area | Specific research assessed | Methods | IDOs and SLOs targeted[b] |
|---|---|---|---|---|---|
| 11. Maize-legume intercropping and rotation, improved maize varieties in Ethiopia, and production diversity (number of food crops produced) in Malawi<br>(*Malawi: Koppmair, Kassie, and Qaim, 2017; Ethiopia: Kassie et al., 2016*) | Nutrition | Production system management | Studies are linked to the International Maize and Wheat Improvement Center's (CIMMYT) ongoing varietal development and conservation agriculture, and to the Sustainable Intensification of Maize-Legume Cropping Systems for Food Security in Eastern and Southern Africa (SIMLESA) project. | Quasi-experimental (fixed-effects multinomial endogenous switching regression) for Ethiopia<br>Quasi-experimental (generalized Poisson regression model) for Malawi | Nutrition (SLO 2): dietary diversity, calorie intake, protein and iron consumption, stunting |
| 12. Water management: irrigated horticulture (*Fanzo, Fishman, and Downs, 2017*) | Nutrition | Production system management | Built on the African Market Garden (AMG) model co-developed by the ICRISAT Sahelian Center from 2001 to 2011 (water, inputs), the PAPSEN-TIPA project (of which the RCT is a part) started in 2015. | Experimental | Adoption<br>Productivity (SLO 1)<br>Nutrition (SLO 2)<br>Women's income and time use (Gender CCT) |
| 13. Drought-resistant rice (BD56) in Bangladesh (*de Janvry et al., 2017*) | RCT | Germplasm enhancement and breeding | BRRI Dhan 56 (BD56) is a drought-tolerant and early-maturing rice variety that was first released in 2011 but has not yet been widely diffused. Development of BD56 is a part of the Stress Tolerant Rice for Africa and South Asia (STRASA) project of IRRI. | Experimental | Productivity, income (SLO 1)<br>Water use efficiency (SLO 3)<br>Farmer capacity for innovation: learning (Capacity development CCT) |
| 14. Maize diversification and intercropping in Nepal (*Mobarak, 2017*) | RCT | Production system management | Study looks at alternative approaches to scaling up maize intercropping practices developed by CIMMYT's Hill Maize Research Project (HMRP) from 1999 to 2010. | Experimental | Adoption<br>Productivity (SLO 1)<br>Extension agent and farmer capacity for innovation: knowledge and learning (Capacity development CCT) |
| 15. Integrated soil fertility management (ISFM) in Kenya (*Laajaj and Macours, 2017*) | RCT | Production system management | Study builds on Tropical Soil Biology and Fertility (TSBF)/CIAT research on ISFM starting in 2001, specifically on seed and fertilizer packages for soya and maize developed through farmer participatory research approaches. | Experimental | Productivity (SLO 1)<br>Farmer capacity for innovation: learning (Capacity development CCT) |
| 16. Soil fertility and input management based on soil analysis recommendations (*Mahajan et al., 2017*) | RCT | Production system management | Study is based on CIMMYT's work on conservation agriculture and soil fertility management (analysis-based input recommendations) in Mexico. | Experimental | Adoption<br>Productivity (SLO 1)<br>Farmer capacity for innovation: knowledge and learning (Capacity development CCT) |
| 17. Water conservation and management: *demi-lunes* (*Aker, 2017*) | RCT | Production system management | ICRISAT has been researching alternative rainwater harvesting techniques (*demi-lunes*, zaï pits) in the region since the 1990s. *Demi-lunes* are a part of ICRISAT's Bioreclamation of Degraded Lands (BDS) system and included in its dissemination programs. | Experimental | Adoption<br>Productivity, input use, profits (SLO 1)<br>Farmer capacity for innovation: learning (Capacity development CCT) |
| 18. Agroforestry program in Kenya (*Hughes et al., 2017*) | UEA | Production system management | ICRAF and VI worked together intensively from the early 1990s to 2004. A process tracing study links ICRAF's research to VI's promotion activities in a range of areas: *Calliandra calothyrsus* and *Sesbania sesban* as protein-rich fodder; both herbaceous and woody improved fallows; alley cropping for erosion control and soil fertility; and participatory tree domestication program. | Quasi-experimental (difference-in-difference combined with propensity score matching) | Adoption<br>Productivity, income (SLO 1)<br>Food security, nutrition, dietary diversity (SLO 2)<br>Tree/vegetation cover, soil erosion, soil organic carbon (SLO 3, Climate change CCT) |

| Innovation | SIAC call window[a] | Research area | Specific research assessed | Methods | IDOs and SLOs targeted[b] |
|---|---|---|---|---|---|
| 19. Water conservation and management: alternate wetting and drying (*Rejesus et al., 2017*) | UEA | Production system management | Approach was developed as part of the Irrigated Rice Research Consortium (IRRC) started in 1997, and the first demonstration trials occurred in the Philippines in 2002. | Experimental | Productivity, income (SLO 1)<br>Water use efficiency (SLO 3) |
| 20. Improved forage grasses (*brachiaria*) in Colombia and Nicaragua (*Labarta et al., 2017b*) | UEA | Production system management | Since the early 1970s, CIAT has promoted the transfer of *brachiaria* grasses. At least four cultivars selected for productivity and soil health benefits were released in Latin America in the early 1980s and 1990s. *Brachiaria* breeding started in CIAT in the late 1990s, and at least 3 hybrids have been released. Since 2002, CIAT has worked in an alliance with the private sector to promote dissemination. | Quasi-experimental (instrumental variable, propensity score matching) | Adoption<br>Productivity (SLO 1)<br>Soil carbon (SLO 3, Climate change CCT) |
| 21. Better forest governance and management in Guinea (*Mills, Nelson, and Achdiawan, 2017*) | UEA | Policy | The Landscape Management of Improved Livelihoods (LAMIL) project operated from 2005 to 2008 and was informed by research on forest management and governance done by CIFOR and ICRAF.LAMIL also built on a prior project in these reserves and two others—the Expanded Natural Resource Management Activity (ENRMA) project between 1999 and 2005—run with limited success | Quasi-experimental (difference-in-difference) | Productivity, Income (SLO 1)<br>Food security (SLO 2)<br>Forest cover (SLO 3) |
| 22. Multi-crop, multi-country (*Gollin, Hansen, and Wingender, 2018*) | Other | Germplasm enhancement and breeding | Modern varieties from CGIAR, 1960–2000 (84 countries) | Macro-level economic impact assessment<br>(difference-in-difference and general equilibrium model) | Income (SLO 1)<br>Health (SLO 2)<br>Land use: Rate of agricultural land expansion (SLO 3) |
| 23. Conservation agriculture in Zimbabwe (*Michler et al., 2017*) | Other | Production system management | In 2004, ICRISAT started providing technical assistance to 10 NGOs promoting conservation agriculture for multiple crops. Specifically, it promoted CA based on hand-hoe prepared planting basins. | Quasi-experimental (instrumental variable) | Productivity (SLO 1) |
| 24. Improved value chain: small- and medium-scale enterprises in the furniture value chain in Indonesia (*Clements, Alwang, and Achdiawan, 2017*) | Other | Policy | The assessment focuses on effects of membership in the APKJ, a producer association started as part of the Jepara furniture value chain project. The research on small- and medium-scale enterprises in the furniture industry was conducted between 2008 and 2013. | Quasi-experimental (propensity score matching) | Income (SLO 1) |
| 25. Sustainable Wetlands Adaptation and Mitigation Project (SWAMP) in Indonesia (*Flores, 2016*) | Other | Policy | CIFOR's SWAMP project ran between 2012 and 2015, but research on wetlands and carbon sequestration dates back to 2008–2009. CIFOR scientists claim a direct influence on the development of Indonesia's forest moratorium enacted in 2011. | Qualitative (outcome assessment, integrated assessment models for benefits using before-after approach) | Deforestation (SLO 3)<br>Carbon sequestration (Climate change CCT) |

**a** LTLS = long-term or large-scale; RCT = experimental studies; UEA = underevaluated areas of research.

**b** CCT = cross-cutting theme.

**Table 3.** Subset of SIAC studies that were accountability-oriented and generated results on SRF outcomes at scale (reported in section 4)

| Study | Data and methods | New findings on adoption and productivity | Main findings associated with SLOs |
|---|---|---|---|
| *Wossen et al., 2017* | Instrumental variable, regression, economic surplus model<br>Cross-sectional survey: 2,500 households covering more than 5,000 plots; 16 states representing 80% of Nigeria's cassava area.<br>**REPRESENTATIVE OF:** Cassava-growing households in Nigeria's main cassava regions. | Adoption: 66.7% adoption.<br>Production/productivity: 82% increase in yield (DNA fingerprinting-based adoption estimates). | SLO 1: 4.7% reduction in poverty; 1.8 million fewer households (8.4% of rural poor cassava producers) in poverty in 2015/16.<br>SLO 2: Positive effects on food security measures: per capita food expenditure increased by 24%, and food access/ availability as reported by adults increased by 18%. |
| *ICARDA, 2017* | Instrumental variable, endogenous switching regression<br>Cross-sectional survey of 1,000 lentil-farming households (out of 644,988) from 52 villages across all 10 western Bangladesh districts in Aug–Nov 2015.<br>DNA fingerprinting (grain samples) from 1,694 lentil plots of the 1,000 sampled households.<br>**REPRESENTATIVE OF:** Lentil-growing households in western Bangladesh. | Adoption: 98% of lentil area was planted to varieties released since 1991, with 82 percent of the area planted to the three most important varieties (BARIs).<br>Productivity: Varieties released after 2005 are associated with yield increases of between 356 and 382 kg/ha (27 percent) for those who adopted. | SLO 1: Adoption was associated with higher net income and higher share of total income from lentil production, with an increase in total production of 25,826 tons, worth US$26 million, annually in recent years. |
| *CIP, 2017*<br>*Myrick, 2016* | Economic surplus model<br>Cross-sectional survey of 615 households located in 41 villages, in an area representative of Yunnan potato production in July–Sep 2015.<br>**REPRESENTATIVE OF:** Potato farmers in Yunnan province, China. | Adoption: 23% of households plant C88.<br>Productivity: 26 percent higher yield per hectare compared with weighted average of all varieties excluding C88. | SLO 1: The estimated present value of benefits from planting C88 in Yunnan ranged from a low of US$2.84 billion to a high of US$3.73 billion. |
| *Benzie and Lind, 2017* | Cross-sectional survey: 99 hatcheries from Bangladesh (about half of the country-level number) and 104 from the Philippines were surveyed) (i.e., fish strains collected and DNA fingerprinted).<br>**REPRESENTATIVE OF:** Hatcheries in Bangladesh and Philippines. | Productivity: GIFT or GIFT-derived strains were the most commonly produced strains in both countries, accounting for almost 53% of production in the sampled hatcheries in Bangladesh and 40% of that in the Philippines. About half the farmed fish production in each of the countries in 2015 came from GIFT or GIFT-derived fish. | |
| *Fishman et al., 2017* | Modern variety diffusion index was constructed using Evenson and Gollin (2003) on MV diffusion and spatially explicit crop area maps for 10 crops, 36 countries. Infant mortality data from the Demographic and Health Surveys for 18,382 villages in 437 administrative regions spread across 36 countries, focusing on children born between 1960 and 2000.<br>**REPRESENTATIVE OF:** 36 countries and 10 specific crop-growing regions | | SLO 2: Observed level of MV diffusion reduced the mortality rate of infants by 3-4 percentage points, which translates into at least 3-5 million infant deaths averted per year by the year 2000.<br>The results also suggest that the diffusion of MVs and broad-based increases in agricultural yields contributed to improvements in infant health. |
| *Labarta et al., 2017a* | Nationally representative adoption study (cross-sectional survey). Bean samples collected for DNA fingerprinting, but analysis wasn't done – study reports only self-reported high-iron beans adoption rate.<br>Cross-sectional survey of 19,575 households across 120 randomly-selected Rwandan villages in season B 2015 to identify bean-growing households, and an in-depth survey of 1,397 households from that sample in 2016 (approximately 12 households per village).<br>**REPRESENTATIVE OF:** Bean-growing households in Rwanda. | Adoption: 21% of bean-producing households growing any high-iron beans, with adopting households only planting an average of 7.7 kg (median 3 kg) of high-iron beans in 2015 season B | |

| Study | Data and methods | New findings on adoption and productivity | Main findings associated with SLOs |
|---|---|---|---|
| *Malawi: Koppmair, Kassie, and Qaim, 2017; Ethiopia: Kassie et al., 2017* | **Malawi**<br>Cross-sectional survey (nationally representative) of 1,482 farm households in 16 districts and 165 villages of Malawi in 2014. Generalized Poisson regression model used to assess associations between intervention and outcomes.<br>**REPRESENTATIVE OF:** Maize-growing farm households in Malawi. | Adoption: Maize-legume intercropping practiced by about 50 percent of households. 81% of households self-report growing improved maize, and 61% report growing legumes. | SLO 2: More diversified farming systems are associated with more diverse diets, but the effects are small. Access to markets (SLO 1) for buying food and selling farm produce, and use of chemical fertilizers are shown to be more important for dietary diversity than diverse farm production. |
| | **Ethiopia**<br>Panel survey had targeted all maize-growing agroecological zones. In 2010, 2,400 farm households operating on 4,354 maize plots and in 2013 follow-up survey 2,289 households operating on 3,907 maize plots were interviewed across 39 districts and 4 regional states.<br>Fixed-effects multinomial switching regression adapted to assess impacts of four "treatments" derived from a combination of cropping system diversification, defined as maize-legume intercropping or rotation, and use of improved maize seeds.<br>**REPRESENTATIVE OF:** Maize-growing households in relevant agroecological zones of Ethiopia. | Adoption: 97% of households in 2010 and 95% of households in 2013 grew maize. About 37% and 13% of maize plots were not treated with both crop system diversification (CSD) and improved maize varieties in 2010 and 2013 respectively. 54% and 65% of maize plots treated with improved seeds in 2010 and 2013 respectively, and 9% and 2% of plots treated with CSD in 2010 and 2013 respectively. Haricot bean is the main legume that is intercropped. Simultaneous adoption of both improved maize varieties (self-reported) and crop diversification (maize-legume intercropping or rotation) was 8% in 2010 and 29% in 2013. | SLO 2: Adoption of cropping system diversification (CSD) and improved maize varieties are associated with improvements in child stunting indicators, per capita consumption of calories, protein, and iron; and a measure of the diversity of food consumed in the household. The greatest impact was achieved when farmers adopted CSD and improved maize varieties jointly rather than individually. |
| *Hughes et al., 2017* | Difference-in-difference combined with propensity score matching (recall baseline + remote-sensing imagery)<br>Survey of 1,450 households belonging to 226 preexisting farmer groups in 60 targeted program villages, and 1,410 households in 206 preexisting farmer groups operating in 61 geospatially matched comparison villages.<br>Systematic sex disaggregation of outcome variables.<br>GPS coordinates collected from surveyed plots of each household's main parcel were used to estimate (using Land Degradation Surveillance Framework) tree cover, vegetation cover, soil organic carbon, and soil erosion prevalence.<br>**REPRESENTATIVE OF:** Vi Agroforestry program beneficiaries in Kenya. | Productivity: Dairy farmers were targeted for production and use of tree/fodder shrubs, and the study examined the yield and input use for this subgroup. About 50% of dairy farmers were using fodder in 2016, an increase of 27% between 2007 and 2016 (10% for the comparison group). Milk yields increased by 0.2 liter a day. | Also reports results on the training conducted under the Vi Agroforestry project. 50% of Vi group members reported being trained and had implemented this training to at least a medium extent in the previous three years.<br>SLO 1: Significant yet variable effects on agroforestry product income and fuelwood access. Statistically significant though modest program effects on asset and consumption expenditure measures (including the primary outcome variable), particularly among households represented by female program participants.<br>SLO 2: Milk yields increased by 0.2 liter a day, but these did not contribute to the minimum dietary diversity for women (MDD-W) impacts.<br>Modest, positive effects for MDD-W were found in program areas, but there was no significant difference between program and comparison groups on the food security measure (months of adequate food provisioning).<br>SLO 3: Program area plots gained an additional 3% tree cover over comparison plots between 2007 and 2016. Vegetative cover was lower in 2016 than in 2007 in both program and comparison areas. No significant difference between program and comparison plots was found for soil erosion or soil organic carbon. |

| Study | Data and methods | New findings on adoption and productivity | Main findings associated with SLOs |
|---|---|---|---|
| *Labarta et al., 2017b* | Planned instrumental variable, propensity score matching but to date only adoption.<br>Nationally representative household surveys.<br>Adoption: 1041 livestock-producing households of 102 different farming communities in the Eastern Plains and Amazon regions of Colombia; 480 households of 80 farming communities in Nicaragua.<br>Expert opinion workshops (Peru, Costa Rica, Honduras).<br>**REPRESENTATIVE OF:** Livestock-producing households in Colombia's Amazon and Eastern Plains regions and Nicaragua. | Adoption: Of the 13.3 million hectares of pastures in the selected countries, about 30% are planted with CIAT-selected *brachiarias*—35% in Colombia and 16% in Nicaragua. The uptake of CIAT *brachiaria* hybrids is only 0.5% of the total area under pastures in the selected countries. | |
| *Mills, Nelson, and Achdiawan, 2017* | Difference-in-difference (recall baseline + remote-sensing imagery)<br>Survey of 240 households representative of 4 LAMIL program reserves, and 240 control households in counterfactual/comparison reserves (4 reserves in total) matched (by GIS expert and local forest/agriculture officers).<br>ASTER imagery, with the occasional use of LANDSAT, was used to interpret land cover changes for three periods (1999–2004 when co-management occurred; 2005–2009 when co-management was revised; and 2010–2014 after the project ended) for the reserves as well as a 5-km buffer area around them.<br>**REPRESENTATIVE OF:** LAMIL program reserves in Guinea. | | SLO 2: Some evidence of higher levels of food security, no evidence of higher income flows from the production of maize, groundnuts, tree plantations or live fencing.<br>SLO 3: Forest reserve land classification analysis (as natural forest or human and agricultural use) using remote-sensing imagery reveals an overall declining trend (higher decline each year starting in 2004) in polygons classified as natural forest, relative to 1999 levels: however, the decline is lower in LAMIL areas by 2%, almost offsetting the overall 2.5% rate of decline. The area of natural forest retained in LAMIL is hence estimated at about 2,400 ha (total LAMIL reserve area), with associated benefits from sequestered carbon estimated between US$7 million and US$14 million, depending on social costs of carbon. |
| *Gollin, Hansen, and Wingender, 2018* | Difference-in-difference and general equilibrium model.<br>Adoption data for 10 crops in 84 countries over 1960–2000—combining newly available Global Agro-ecological Zone (GAEZ) data on agro-climatically attainable crop yields under local climatic conditions, and the spatial/temporal variations in diffusion patterns of HYVs.<br>**REPRESENTATIVE OF:** 84 countries and 10 specific crop-growing regions | | SLO 1: Productivity increases resulted in a 139 percentage points (pp) contribution to growth. 60% of the GDP per capita effect can be attributed to productivity increases, and about 40% comes from movement of agricultural workers into more productive occupations.<br>SLO 2: The Green Revolution reduced fertility, and the reduction was only partly offset by decreasing mortality rates. The net effect on population growth was therefore negative. 10 percentage-point increase in HYV adoption was estimated to increase life expectancy by 1.34 percent.<br>SLO 3: Increases in land area under HYV tended to reduce total land under cultivation, consistent with "Borlaug hypothesis". |
| *Michler et al., 2017*<br>*and*<br>*Michler et al., 2018* | Instrumental variable.<br>Four years of plot-level panel data covering 5 crops (cowpea, groundnut, maize, millet, sorghum,) for 2007–2015; 4,217 plots from 730 households across 45 wards of Zimbabwe; rainfall shocks calculated from satellite imagery from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS).<br>**REPRESENTATIVE OF:** Farm households in Zimbabwe. | Productivity: No or possible negative effect on yields in average rainfall years for all crops (cowpea, maize, sorghum, millet, groundnut). Positive effect on yields, especially for maize, in years of very low rainfall. | SLO 1: Gross revenues to CA are less than gross revenues to traditional practice over a large portion of the rainfall distribution. Only at the far ends of the distribution, particularly in very wet years, does CA produce more gross revenues. |

| Study | Data and methods | New findings on adoption and productivity | Main findings associated with SLOs |
|---|---|---|---|
| *Clements, Alwang and Achdiawan, 2017* | Propensity score matching, baseline constructed on recall.<br>Propensity score matching was used to compare outcome variables and changes in these variables of association members and matched nonmembers.<br>**REPRESENTATIVE OF:** Furniture producer associations in Indonesia. | Adoption: Membership in producer associations has not grown over time. This might be because membership had no significant effect on profit levels even 6 years after the project commenced. | SLO 1: Membership had no significant effect on firm profit levels. |
| *Flores, 2016* | Qualitative: outcome assessment, integrated assessment models for benefits using before-after approach.<br>**REPRESENTATIVE OF:** Indonesian forested wetlands. | | SLO 3: The 2011 moratorium was not effective in protecting forests in non-designated areas expected to gain protection under the moratorium, and deforestation increased as compared to pre-2011 trends. Similarly, concession areas (i.e., land licensed to oil palm, timber, or logging plantations) experienced a much higher rate of forest loss after the 2011 moratorium. However, protected areas had a lower rate of deforestation. |

**Table 4.** Subset of SIAC studies that were learning-oriented (reported in section 5)

| Study | Data and methods | Main findings on SLOs |
|---|---|---|
| *Glennerster and Suri, 2018* | Randomized control trial: multiple treatments, community-level. At the community level (36 communities in pilot phase, 293 for full trial), pricing (free, half price, full price) and training on NERICA-3 were randomized; at the household level, offer of sale was randomized.<br>Survey of 2,907 farm households in 6 districts of Sierra Leone. | Adoption: 97% adoption of NERICA-3 in the group offered seeds free. Lower take-up (62%, 20%) for households that were offered seeds at full price and half price, respectively.<br>Productivity: 23% increase (45 kg per acre) in yields relative to control, but only for those offered training and free seeds.<br>SLO 2: Positive impacts on weight-for-height and body-mass-for-age Z-scores for households offered training and free seeds; stronger effect in the second season than the first. |
| *de Janvry et al., 2017* | Randomized control trial, village-level: 3 treatment arms. 192 villages given BD-56 seed mini-kits, and 64 villages in control group given analogous variety BD-51 seed mini-kits. Second level of randomization by types of farmers in 192 villages: in each set of 64 villages, the 5 mini-kit recipients were either randomly drawn farmers, largest farmers, or farmers selected by local extension officer.<br>Survey of 1,795 farmers across 256 villages of the Rajsahi Region of Bangladesh, where BD-56 is suitable for cultivation. | Adoption: Take-up of 5-kg seed mini-kits was 64%. Take-up was higher where largest farmers were provided with mini-kits and the seeds were cultivated next to a variety of their choosing. No difference in take-up between BD-56 and BD-51 recipients.<br>Productivity: BD-56 was harvested 25 days earlier than BD-51 with a 43% lower yield. BD-56 plots (*aman* or wet season) were 28 percentage points more likely to be cultivated with a *rabi* or winter crop rather than leave the plot fallow. Planting of a *rabi* crop did not alter the likelihood of a *boro* or dry/summer crop. 52% of farmers cultivated a *rabi* crop (such as mustard, potatoes, pulses).<br>SLO 1: Farmers who cultivated BD-56 and a rabi crop had a 16% higher annual crop profit than control farmers.<br>SLO 3: Slightly fewer watering days (0.5) for BD-56 in the first main rice season. |
| *Mobarak, 2017* | Randomized control trial, ward-level: 9 treatment arms. 48 control wards and 120 treatment wards chosen randomly from 40 agricultural service center (ASC) areas/4,670 wards across 10 districts in Nepal. 120 treatment wards were randomly assigned to one of three types of individuals to disseminate information on maize intercropping with tomato, ginger, or French beans—agricultural extension workers (AEWs)–25 wards, lead farmers (LFs)–50 wards, and peer farmers (PFs)–45 wards. The three types were further randomized by incentives—performance-based payments, flat payments, and no incentives. Survey of 15 households per ward—2,520 households. | Adoption: Increased by 15–30 percentage points across all three types of communicators. At the end of two years, across all treatments, around 20 percentage adoption. Incentives increased adoption by 15–20% over control, but "no incentive" wards performed better than incentive wards.<br>Productivity and SLO 1: No significant impacts on yields or measures of household welfare were found. |

| Study | Data and methods | Main findings on SLOs |
|---|---|---|
| *Laajaj and Macours, 2017* | Randomized control trial, multiple treatments: Across 96 villages, 10 farmers in each village first identified plots they would dedicate to research trials. Among the 10 farmers in each village, 5 were randomly selected and 5 were specifically selected as promising farmers for trials by the community. Half the villages were then assigned to control. Within treatment villages, the trials started in 24 villages in the long rain season of 2014; for the other 24, trials started in the short rain season of 2014. Agronomic farmers from IITA then worked with farmers in treatment group to implement the experimental protocol: each plot was randomly divided into a control subplot without inputs and 5 treatment plots with different combination of inputs (seeds and fertilizers) for soya and maize, and inputs were applied for three consecutive seasons. | Productivity: With optimal fertilizer package, maize yields rose by 30–200% and soya yields rose by 50–150%, with the range explained by heterogeneity (location and varieties used). |
| *Mahajan et al., 2017* | Randomized control trial, household-level: 4 treatment arms, encouragement design. Agricultural extension services cross-randomized with soil analysis and recommendations (individualized, averaged), and grants for inputs (none, flexible, and inflexible).<br>Program widely advertised in Tlaxcala state, Mexico, during 34 promotional meetings and among the 1,299 farmers owning <15 hectares of land; 981 farmers randomized into 4 treatment arms and a control arm. | Adoption: In treatments T1 (individualized plot-level soil analysis and recommendations, inflexible in-kind subsidies, extension), T2 (average village-level soil analysis and recommendations, inflexible in-kind subsidies, extension), and T3 (average village-level soil analysis and recommendations, flexible in-kind subsidies, extension), the take-up rate was more than 75%. In T4 (average village-level soil analysis and recommendations, no subsidies, extension), the take-up rate was only 7%.<br>Productivity: Farmers across treatments continued to use practices learnt (in 2015) in the following year, especially T3 farmers. Farmers were more likely to use fertilizer with machinery at sowing, herbicide one week after sowing, and topdressing. This learning process did not translate into yield increases in 2016. |
| *Aker, 2017* | Pilot randomized control trial, village-level: 3 treatment arms. Training on *demi-lunes* combined with one of the finance options (no grant, conditional cash transfer, unconditional cash transfer). 30 villages were assigned to 1 of the 3 treatments.<br>Survey of 750 farmers from 30 villages (25 in each village) in the Dosso region of Niger. | Adoption: 85% of households that received unconditional cash transfer and training (UCT + T) or training only (T) constructed *demi-lunes*, whereas households that received conditional cash transfer and training (CCT + T) were 13 percentage points less likely to construct *demi-lunes*. *Demi-lune* recommendations (250 per hectare) were not followed, but those who constructed *demi-lunes* followed spacing norms, and qualitative follow-up indicates an increase in number constructed every year. Largest number of *demi-lunes* constructed by farmers in UCT villages.<br>Productivity: 20 kg more millet yield for households in UCT + T villages compared to households in T or CCT villages. Adoption of *demi-lunes* was correlated with changes in input use. Households in UCT + T villages were 18 percentage points less likely to use inorganic fertilizer than those in CCT + T villages but more likely to use organic fertilizer (manure).<br>SLO 1: Adoption was associated with other positive improvements in well-being—households in UCT +T villages were 12–14 percentage points more likely to own a motorcycle and cart than those in CCT + T or T only villages. |
| Rejesus et al., 2017 | Randomized control trial: one treatment (alternate wetting and drying [AWD]), cluster randomization<br>Only 92 valid turnout service area groups (TSAGS) out of 280 TSAGS from the Bicol RIIS were selected for the study. Bicol is the sixth major rice-producing region in the Philippines.<br>Survey of 840 farmers from 42 randomly drawn TSAGs, out of 92 total, in May–June 2016—sample equally distributed between treatment and control. | Adoption: Official adoption estimates obtained from National Irrigation Administration (NIA) for three regions indicate 84,784 hectares over which AWD is practiced, or approximately 60,559 farmers (assuming an average farm size of 1.4 hectares per farmer). These adoption figures roughly correspond to 2% of the 4.5 million ha of rice area in the country.<br>Productivity: No statistically significant impacts on yields or change in management (size of rice parcel, irrigation frequency).<br>SLO 1: No impacts on self-reported gross income.<br>SLO 3: No impacts on the number of days main rice parcel is without water. |

Independent
Science and
Partnership
Council

CGIAR Independent Science & Partnership Council (ISPC) Secretariat
c/o FAO, Viale delle Terme di Caracalla 00153 Rome, Italy
t: +39 06 570 52103 - e: ISPC-Secretariat@fao.org
http://ispc.cgiar.org