# Maximizing Impact from Agricultural Research: Potential of the IAR4D Concept

FARA

European Commission

IFAD

CGIAR

# Maximizing Impact from Agricultural Research: Potential of the IAR4D Concept

A. A. Adekunle, A.B. Ayanwale, A.O. Fatunbi,
A. Agumya, F. Kwesiga and M.P. Jones

**Forum for Agricultural Research in Africa**
12 Anmeda Street, Roman Ridge,
PMB CT 173, Accra, Ghana
2013

**Citation:**

Adekunle. A.A., A.B. Ayanwale, A.O. Fatunbi, A. Agumya, F. Kwesiga and M.P. Jones (2013). *Maximizing Impact from Agricultural Research: Potential of the IAR4D Concept.* Forum for Agricultural Research in Africa (FARA), Accra, Ghana.

FARA encourages fair use of this material. Proper citation is requested.

**Forum for Agricultural Research in Africa (FARA)**
12 Anmeda Street, Roman Ridge
PMB CT 173, Accra, Ghana
Tel: +233 302 772823 / 302 779421
Fax: +233 302 773676
Email: info@fara-africa.org
Website: www.fara-africa.org

ISBN 978-9988-8438-2-8 (print)
ISBN 978-9988-8438-0-9 (pdf)

Editing and design: www.bluepencil.in
Print: www.pragati.com

# Contents


| | |
|---|---|
| **Acknowledgements** | 1 |
| **Executive summary** | 3 |
| Does the IAR4D work as a concept? | 4 |
| Does the IAR4D deliver more benefits than the conventional R&D methods? | 4 |
| Can the IAR4D be scaled up and out beyond the current area of operation? | 5 |
| **Chapter 1: Introduction** | 6 |
| Objectives of the SSA CP | 8 |
| Expected outputs | 8 |
| **Chapter 2**: **Methodological Framework** | 10 |
| Study area | 11 |
| IAR4D impact pathway | 11 |
| Research questions and hypotheses | 13 |
| Analytical approach | 14 |
| Evaluation design | 15 |
| Characterisation of treatment and counterfactual sites | 16 |
| Census of district/local government area/commune sites and characteristics | 18 |
| Sampling method | 18 |
| Sample sizes at IP, Task Force, PLS and SSA CP scales | 19 |
| Sample selection | 20 |
| Baseline surveys for IP and community level characteristics | 20 |
| Baseline survey for household and village community characteristics | 20 |
| Evaluation surveys | 21 |
| Data analysis | 21 |
| **Chapter 3: Results and Discussion** | 26 |
| Impact of IAR4D on household income | 27 |
| Estimation results of propensity scores | 27 |
| **Chapter 4: Conclusions and Policy Implications** | 34 |
| Does the IAR4D work as a concept? | 35 |
| Does the IAR4D deliver more benefits than the conventional R&D methods? | 35 |
| Can the IAR4D be scaled up and out beyond the current area of operation? | 36 |

**Annex 1.** Methodological Framework for Impact Evaluation of the Sub-Saharan Africa Challenge Programme (SSA CP) 39

**References** 59

**Acronyms and abbreviations** 62

**List of Tables**

Table 1. Research questions and hypotheses 13
Table 2. Sample sizes at IP, task force, PLS and programme scales 19
Table 3. Variables used to compute propensity scores and their expected signs 24
Table 4. Impact of IAR4D on household income across types of respondents 28
Table 5. Probit regression of IAR4D participation (matched observations) 30
Table 6. Impact of IAR4D on income distribution 32
Table A1. Summary of the possible approaches to impact evaluation. 41

**List of Figures**

Figure 1. IAR4D impact pathway 13
Figure 2. Illustration of stratification of a PLS by four development domains 17
Figure A1. The levels at which SSA CP data will be collected and pooled 49

# Acknowledgements

This book documents a study on the concept of Integrated Agricultural Research for Development (IAR4D) that was developed by the Forum for Agricultural Research in Africa (FARA). The IAR4D concept forms the basis for the Sub-Saharan Africa Challenge Programme (SSA CP), which is the only CGIAR Challenge Programme limited geographically to a particular region of the world. The focus of SSA CP is to facilitate substantially greater impact from agricultural research for development (ARD), leading to improved rural livelihoods, increased food security, and sustainable natural resource management throughout sub-Saharan Africa. The SSA CP aims to achieve this aim by developing and implementing the IAR4D approach that overcomes the shortcomings of traditional approaches used in ARD. The programme had been challenged in the research phase, to conduct activities that will lead to a proof of the effectiveness and superiority of the IAR4D concept in meeting developmental goals.

The implementation of SSA CP necessitated extensive partnership arrangements, almost unprecedented in the history of the CGIAR. The implementation was carried out in the three sub-regions that constituted sub-Saharan Africa – West Africa, Eastern and Central Africa, and Southern Africa. More than 80 institutions were involved in the implementation of the programme. While 55% of these institutions were purely research based, the other 45% were civil society organizations (NGOs, private sector organizations, farmers' organizations and community-based organizations). The implementation phase benefited from the inputs of more than 243 researchers across the globe.

Therefore, FARA and the key authors of this book would like to acknowledge the contributions of all the stakeholders of African agriculture that are the constituents of FARA, but more specifically the following categories of stakeholders.

We would like to acknowledge the CGIAR's role in accepting the Programme as a Challenge Programme, and also acknowledge the development partners who funded the SSA CP programme from its inception. The European Commission (EC) and the International Fund for Agricultural Development (IFAD) are especially acknowledged for their unflinching funding support to the Programme at all times. The Government of Italy, the Government of Netherlands, The Government of Norway, the Department for International Development, United Kingdom (DFID UK) and the United States Agency for International Development (USAID) are also similarly acknowledged.

We categorically acknowledge the contributions of Sub-Regional Organizations for agricultural research and development in Africa. We particularly thank the Conseil Ouest et Centre Africain pour la Recherche et le Développement Agricoles/West and Central African Council for Agricultural Research and Development (CORAF/WECARD) and the Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA), for coordinating the West Africa and the East Africa pilot learning sites, respectively. We also thank the Southern African Development Community and the Centre for Coordination of Agricultural Research and

Development for Southern Africa (CCARDESA) for coordinating the third pilot learning site in southern Africa.

Contributions of the following international research institutions as key partners are also recognized: International Institute of Tropical Agriculture (IITA), Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT), Biodiversity International, AfricaRice, International Food Policy Research Institute (IFPRI), International Livestock Research Institute (ILRI), International Crops Research Center for the Semi-Arid Tropics (ICRISAT), and International Fertilizer Development Center (IFDC). Numerous national research institutions also participated as key partners in the implementation of the Programme at the task force level. They include the Rwanda Agricultural Board (RAB), formerly called the Institut National de Sciences Agronomiques du Rwanda (ISAR), the Institut National de Recherches Agronomiques du Niger (INRAN, Niger), Makerere University, Uganda, and Ahmadu Bello University, Zaria, Nigeria.

Other collaborating institutions on the innovation platform (IP) are very much appreciated for their activities on the different aspects of the innovation platform. They include Asian Vegetable Research and Development Centre (AVRDC); Department of Agricultural Research Services (DARS) Malawi; Instituto de Investigação Agraria de Moçambique (IIAM); Department of Agricultural Extension, University of Zimbabwe; World Vision International, Malawi; Olivine PVT, Zimbabwe; Katsina State Ministry of Agriculture, Nigeria; Jubaili Agro-chemical suppliers, Kano, Nigeria; Institute of Agricultural Research, ABU, Zaria, Nigria; National Animal Production Research Institute, Nigeria; Bayero University, Kano, Nigeria; Kano State Agricultural and Rural Development Programme; Women Farmers Association of Nigeria (WOFAN); Katsina State Agricultural and Rural Development Programme (KTARDA)/Ministry of Agriculture; National Seed Council of Nigeria; Project Seeds Nigeria Limited, Kano, Nigeria; Grand Cereals Company, Jos, Nigeria; Modern Universal Foods, Kano, Nigeria; Convenient Foods, Kano, Nigeria; and Golden Fertilizer, Nigeria.

Listing the Scientists for acknowledgement is a difficult task indeed, because many people contributed to the work. However, FARA would like to mention the following scientists for their contribution to the research on IAR4D at different stages: Prof Alphonse Emechebe (CORAF/ WECARD); Dr Robin Buruchara (CIAT); Drs Kamara Alpha, Joseph Rusike, Tairou Abdulahi and Mrs Yemisi Micha (IITA); Dr Nelson Mango and Prof Wanjiku Chiuri (CIAT); Dr Jojo Baidu-Forson (Bioversity International); Prof Uzo Mokwunye (Independent Consultant); Prof Moses Tenywa (Makerere University, Uganda); Dr Josephat Mugabo (RAB, Rwanda); Dr Nouri K. Maman (INRAN, Niger); Drs Constance Dangbenon and Abdulahi Mando (IFDC); Professors Ben Ahamed, Ishiaku Amapu, Chris Daudu and Drs Adunni Sanni and Damisa Maiyaki (IAR/ ABU, Zaria); Prof Luke Olarinde (Ladoke Akintola University, Ogbomoso); Drs Shepherd Siziba and Kefasi Nyikahadzoi (University of Zimbabwe); Dr Judith Odoul, Mr Ben Fungo and Mr Peter Muraya (ICRAF, Nairobi); Dr Joachim Binam (ICRISAT, Mali); Prof Adeolu Ayanwale (Obafemi Awolowo University, Ile-Ife); Prof Sagary Nokoe (University of Development Studies, Ghana); Dr Aliou Diagne (Africa Rice); Dr Epraim Nkoya (IFPRI, Washington, D.C.); Dr Olufemi Oladunni (ARMTI, Ilorin); Dr Gloria Essilfie (University of Ghana); Prof Ewen Bulte (University of Wageningen) who worked with a host of others to ensure we accomplished the task.

# Executive Summary

Most of the 800 million people that inhabit sub-Saharan Africa (SSA) do not only live in rural areas, but also have agriculture as their major livelihood activity and major employer of labour, providing over 60% of full-time employment and generating 27% of gross domestic product (FAO 2009; World Bank 2008). In essence, the region's overall economic performance is inextricably linked to the performance of its agricultural sector. But agriculture in SSA has underperformed, and food insecurity is rampant. Part of the reason has been the limited impact of past agricultural research on intended beneficiaries. The use of traditional approaches for research and development is widely blamed for the poor performance of Africa's agricultural sector, as it resulted in low adoption rates of technologies, poor linkages among agricultural value-chain actors, and the chronic nonprofitability of farm enterprises in SSA.

The acknowledged poor performance of traditional ARD approaches led the Forum for Agricultural Research in Africa (FARA) to suggest the Integrated Agricultural Research for Development (IAR4D) as an approach to enable agricultural research play more effective roles in catalysing development, by embracing a broader system of agricultural innovation that will facilitate interaction and enhance the flow of knowledge among all key actors in agricultural systems and value chains. FARA developed a programme around the concept of IAR4D and this was accepted by the CGIAR as the Sub-Saharan Africa Challenge Programme (SSA CP).

Through the SSA CP, IAR4D is being implemented in three Pilot Learning Sites (PLS) across the continent with the central aim of reversing the underperformance of agricultural research in Africa. Specifically, this is being done by developing, testing (proving whether it works) and scaling out/up an approach for conducting agricultural research for development in Africa, which overcomes the shortcomings of conventional approaches. Each PLS defines the domain within which the project's research sites are sampled. This study is focused on the three PLSs that made up the SSA CP.

With clearly defined outputs in mind, the SSA CP was mandated by the Science Council (SC) of the CGIAR to commence a proof of the concept research phase, with the aim of answering three vital questions as to the relevance and effectiveness of IAR4D in delivering developmental benefits and its relative performance when compared with conventional approaches in promoting impact.

These questions were as follows:

- Does the IAR4D work?
- Does the IAR4D deliver more benefits than the conventional R&D if given the same environment and resources? and
- Can the IAR4D be scaled up and out?

These questions were the motivation for this report. The report made use of data collected from baseline and midline surveys, organized using the quasi-experimental approach, and two sets of counterfactuals, namely, the conventional or traditional ARD, and the clean sites where it was assumed there was no ARD at least two years prior to the commencement of the IAR4D experiment.

Using propensity score matching (PSM) and double-difference methods (DDM) to control for project placement and self-selection biases, we found that IAR4D improved the household assets of the participants, as well as encouraged participation in research and facilitated the adoption of research outputs.

The results of the probit regression show that the participants in the IAR4D would most likely be young married farmers with a small family size. However, there is a need to consciously encourage participation of farmers from Malawi, Mozambique, or Zimbabwe. Participants in the conventional module are mostly farmers with some level of productive assets, with those from Zimbabwe needing conscious encouragement, while those in the clean sites are female farmers without productive assets, with those from Mozambique needing conscious encouragement. This result suggests that the IAR4D intervention focused on married youths, who are the more vulnerable groups in sub-Saharan Africa.

## Does the IAR4D work as a concept?

The answer to this question came from the homogenous result of the impact analysis. The answer is yes; the IAR4D works and impacts positively on the lives of the beneficiaries to the tune of $1362.72 per participant per year. This amount translates to a daily income of $3.73, which is more than three times the World Bank's poverty line of $1.00 a day. This is better appreciated when the baseline conditions are considered.

## Does the IAR4D deliver more benefits than the conventional R&D methods?

With the use of matching methods as well as the PSM and double difference approach, we can safely conclude from the results that the IAR4D delivers more benefits than the conventional R&D method. The results, while showing the positive impact for the IAR4D, reveal that under the same conditions, the conventional and the clean do not impact consistently and positively on the non-beneficiaries.

The analyses also show that the IAR4D impacts positively on women's income (326%), food security (324%) and wealth distribution (5%). Clearly, the income of 4,656 women was improved, while 6,504 people were able to cross the poverty line as a result of their participation in the programme. With age-group disaggregation, 7,144 young farmers had their income improved. These results were consistently robust and reliable.

Household incomes improved substantially more for the IAR4D participants than for non-beneficiaries in both conventional and clean sites, with an average increase in real incomes resulting from participation of about 232%. The observed household income is not only better than the conventional and clean sites, but also well above the achievement of similar projects in the continent. For instance, the World Bank sponsored Fadama II project in Nigeria (which won the Bank's Regional Excellence Award) had an income impact rate of about 60 percent, a feat achieved after 6 years of operation.

## Can the IAR4D be scaled up and out beyond the current area of operation?

The results of the ex-ante analysis, in line with the impact assessment analysis, suggest that the concept can be successfully scaled up and out, with potentially multiple positive impact on the beneficiaries.

The findings indicated that estimated benefits will be influenced by adoption to IAR4D, rather than changes in research and extension costs. Nevertheless, the estimates indicate that the production of all the commodities will be profitable under the IAR4D approach. The results were consistent with earlier economic analyses, which showed that IAR4D was more productive, profitable and acceptable to farmers than the conventional Research for Development (R&D) approach.

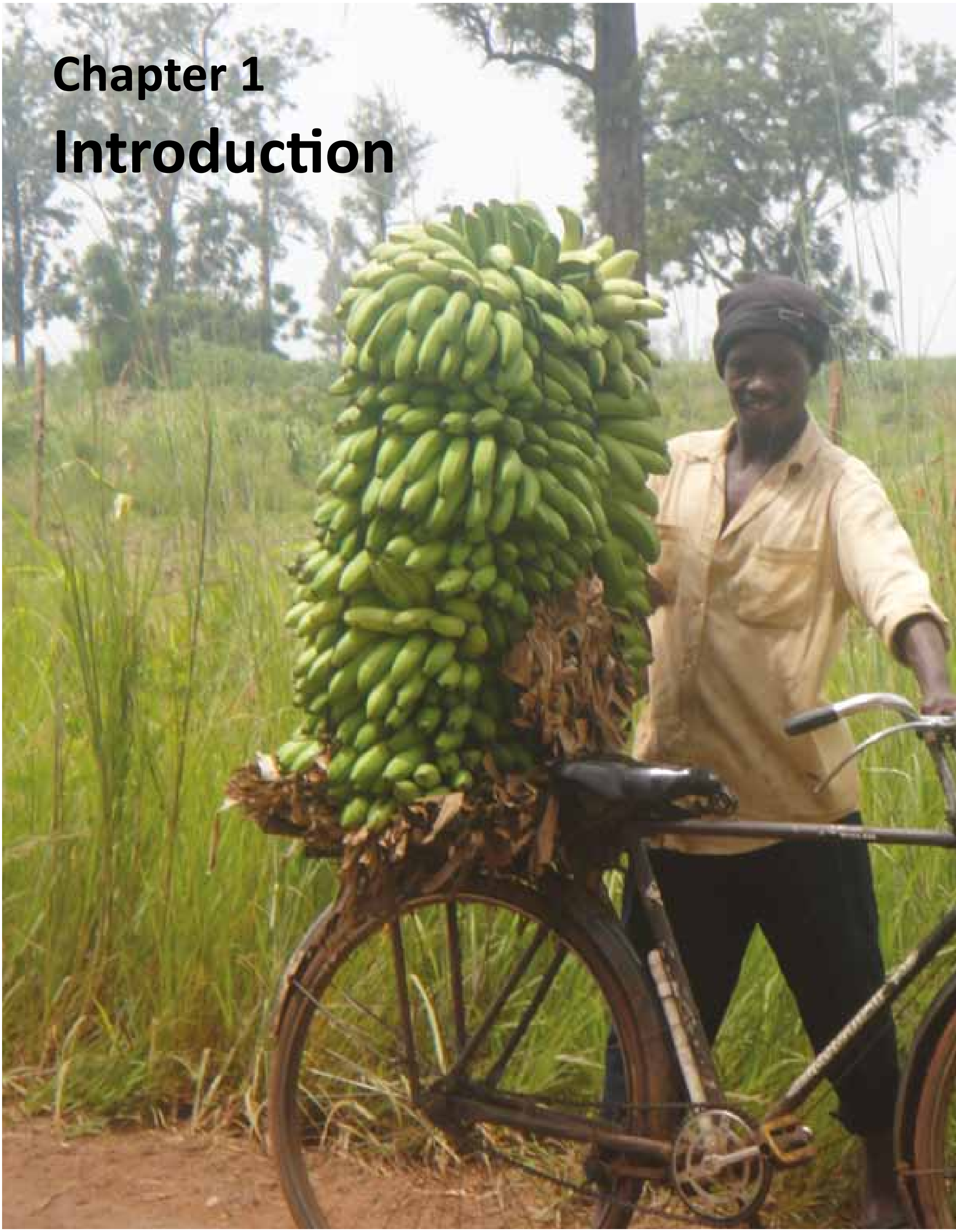
The ex-ante analysis of the three PLS (Ayanwale et.al. 2011) had confirmed that the projected benefits of IAR4D not only surpassed the costs of investments, but they were also superior to both the conventional and clean modes. Furthermore, the benefits derivable varied by task forces (agro-ecological zones), with the Zimbabwe-Mozambique-Malawi (ZMM) PLS showing the least quantum of benefits of the three. This could be owing to the politico-economic situation of the countries.

The project had a bigger impact on the poorest beneficiaries and could have much greater impact in the future because of the lagged effect of the productive asset acquisition. Thus, a follow-up study is needed to capture the longer-term effects of productive assets and other changes that farmers experienced as a result of participation in the IAR4D. This study was conducted at an early stage of the project and does not capture its time-lagged impacts, especially the long-term benefits of productive asset acquisition and rural infrastructure development.

Key issues that need to be addressed in scaling up this success story include, among others, better targeting of poor and vulnerable groups, especially women; finding sustainable methods of promoting development of rural financial services; and the conscious inclusion of capacity building of IAR4D beneficiaries in efficient management of productive assets.

As regards appropriate targeting, it may be recalled that over the first 2 years that the project operated, the Gini coefficient of income for beneficiaries decreased by about 6%, compared with an increase for other categories of non-beneficiaries. This suggested that the project contributed to a reduction in income inequality, most probably through targeting the poor and vulnerable groups. Consistent with this, the project also succeeded in raising the value of productive assets of the poorest tercile more significantly than for the other terciles. The non-significance of the impact on income for the other two terciles suggested appropriate targeting of the poor and vulnerable groups.

# Chapter 1
# Introduction



The sustainable livelihoods of many African people depend directly on their ability to produce and market agricultural products. Consequently, agricultural growth in sub-Saharan Africa remains fundamental for ensuring poverty reduction and food security. It has been realized that without urgent revitalization of the agricultural sector, the Millennium Development Goals (MDGs) to halve poverty and hunger, as well as ensure environmental sustainability, by 2015 will be difficult to meet. As a result, substantial investments have been made in agricultural research and innovation. However, it has been observed over time that the impact of some of the investments has not gone beyond the immediate localities of the research environment, while some efforts have resulted in outright failure.

Consequently, extensive consultations between 2002 and 2004 led to the formulation of the Sub-Saharan Africa Challenge Programme (SSA CP), following the discovery that the principal shortcoming of African agricultural research and development (ARD) has been its failure to achieve impact beyond the localities in which studies had been conducted and the accumulation of so-called 'improved technologies' on research shelves rather than in farmers' fields. Therefore, the SSA CP concluded that for agricultural research to play a more effective role in catalysing development, it should embrace a broader system of agricultural innovation that will facilitate interaction and enhance the flow of knowledge among all key actors in agricultural systems and value chains. FARA has called this systemic and innovation-focused approach to agricultural research as *Integrated Agricultural Research for Development* (IAR4D).

IAR4D seeks to transform the organizational architecture of R&D actors from a linear configuration (research→ dissemination→ adoption) to a network configuration, comprising all actors in the agricultural ***Innovation Sphere***, which basically includes the players in the value chains for commodities, or the players in the value web for systems, as the case may be, but also includes other players who complement these players in the process of putting research to use for economic benefits (the innovation system). The network configuration facilitates timely interaction and learning and aims at generating innovations (rather than research products per se). The innovation in this concept refers to the activities and processes associated with the generation, product distribution, adaptation and use of new technical and institutional/ organizational knowledge. It adds value to products of research to catalyse the achievement of developmental impact.

## Objectives of the SSA CP

The objectives of SSA CP are to facilitate substantially greater impact from agricultural research for development, leading to improved rural livelihoods, increased food security and sustainable natural resource management throughout SSA. The SSA CP is being implemented in three Pilot Learning Sites (PLS) across the continent. By applying IAR4D, the SSA CP aims to reverse the underperformance of agricultural research in Africa by developing, testing (proving whether it works) and scaling out/up an approach for conducting agricultural research for development in Africa, which overcomes the shortcomings of conventional approaches. Each PLS defines the domain within which the project's research sites are sampled. This report is focused on *the three* PLS that made up the SSA CP.

## Expected Outputs

The expected outputs are as follows:

***Output 1:*** Principles, procedures and best practices for implementing IAR4D to generate technological, market, institutional, policy, gender and new product innovations, appropriate to the needs and capabilities of communities in the three PLS.

***Output 2:*** IAR4D-derived technological, market, policy and gender-sensitive innovations and capabilities for sustainably increasing agricultural productivity, value addition and access to agricultural markets by communities in the three PLS.

***Output 3:*** An evaluation of the effect and cost effectiveness of IAR4D on developmental impact (relative to conventional ARD approaches) and the replicability of IAR4D in the various contexts of the three PLS. This will provide empirical proof that IAR4D works in such contexts and is superior to conventional approaches in terms of the benefits it delivers against the costs it entails. The evidence will provide a rationale for the reform of African ARD to reverse the decline in its impact and to increase the likelihood of achievement of the MDGs pertaining to poverty, hunger, empowerment of women and environmental sustainability.

With these outputs in mind, the SSA CP was mandated by the Science Council (SC) of the CGIAR to commence on a "proof of the concept" research phase, with the aim of answering three vital questions as to the effectiveness of IAR4D and its relative performance in the delivery of developmental impact when compared with conventional approaches.

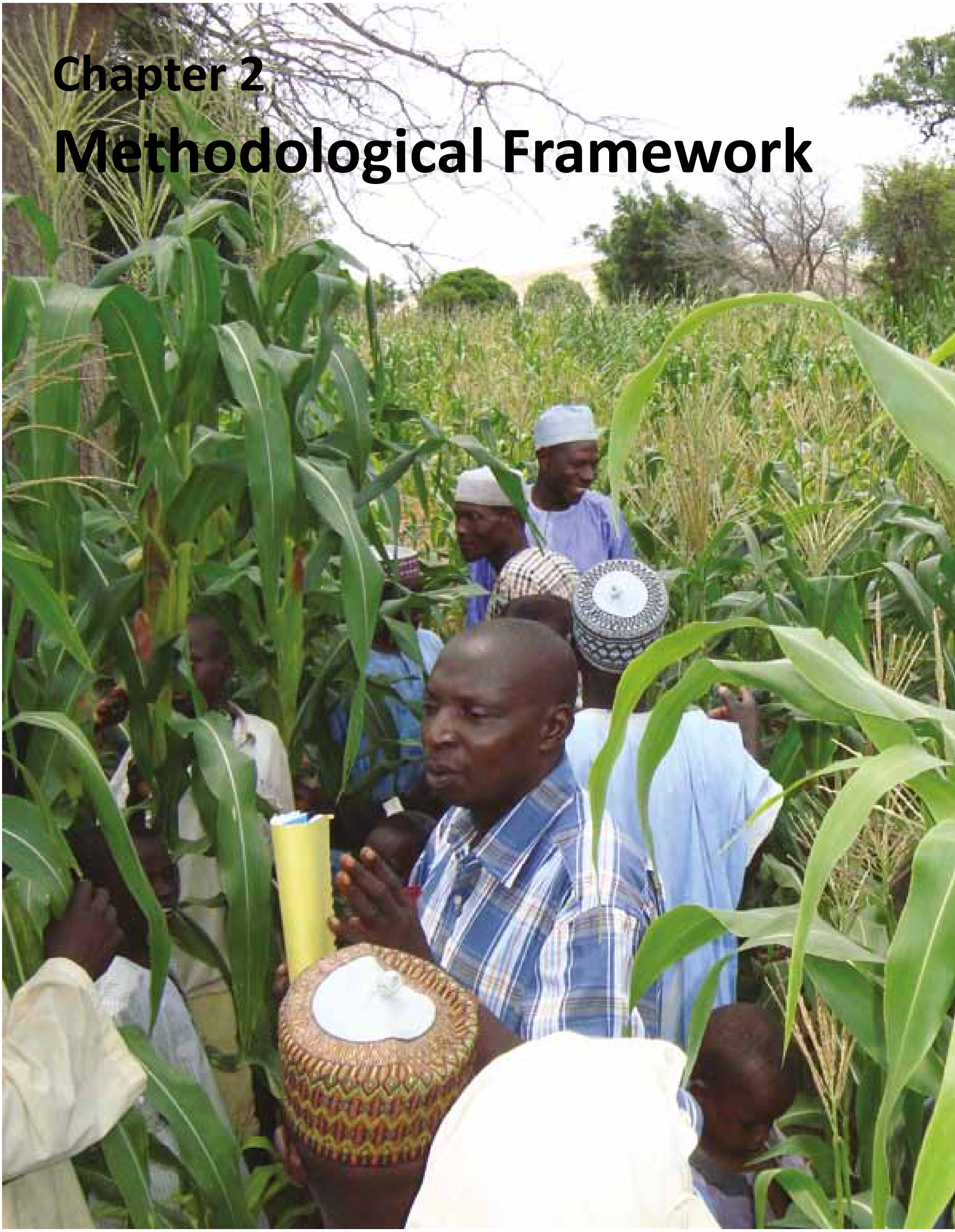
Three key questions were formulated by the SC for the SSA CP to pursue, in its effort to illuminate and provide evidence of the IAR4D concept and its benefits. These questions are as follows:

1. *Does the IAR4D work?*
2. *Does the IAR4D deliver more benefits than the conventional R&D if given the same environment and resources? and*
3. *Can the IAR4D be scaled up and out?*

These questions form the kernel of this report. The report provides answers to these questions to establish the effectiveness of the IAR4D concept, as well as the ability of the concept to deliver more benefits than the conventional concepts.

The rest of this report is divided into three Chapters and an Annex. Chapter 2 relates to the Methodological Framework for this study, Chapter 3 summarizes the Results and Discussion, and Chapter 4 presents the Conclusions and Policy Implications. Given the importance of the methodological framework for researchers in the sub-Saharan African region, it is presented in greater detail in Annex I. Consequently, *References* related to both the main text and Annex are combined and presented at the end.

# Chapter 2
# Methodological Framework



The methodological framework is of paramount importance in a study of this kind. Therefore, this chapter presents its various aspects, such as the study area, the impact pathway envisaged, the research questions and hypotheses, the analytical approach employed, and the evaluation design. The evaluation design, in turn, includes site characterisation methods, local census and sampling methods, sample size, sample selection, baseline surveys, evaluation surveys, and data analysis.

Given that methodological issues are of vital concern to researchers even in evaluating the information presented in this report and in sharing it with others, the Methodological Framework is presented in greater detail in Annex I.

## Study Area

The SSA CP is being implemented in three Pilot Learning Sites (PLS) across the African continent. By applying the IAR4D, SSA CP aims to reverse the underperformance of agricultural research in Africa by developing, testing (proving whether it works) and scaling out/up an approach for conducting agricultural research for development in Africa, which overcomes the shortcomings of conventional approaches. Each PLS defines the domain within which the project's research sites are sampled.

## IAR4D Impact Pathway

Based on the SSA CP's research plan and programme for impact assessment (FARA 2009), the point of departure of IAR4D from traditional ARD approaches lies in how innovations are generated. While traditional ARD approaches *exogenously* bring innovations into the system, IAR4D instead establishes an institutional innovation—the Innovation Platform—which, in turn, *endogenously* generates the innovations (technological, market, institutional and policy). For a summary of the research-to-impact pathway used to hypothesize the causal relationships between research inputs and the research outputs (i.e., the Innovation Platform), institutional innovation and its results (knowledge increase, behavioural change, and innovations at the interfaces of processes driving productivity, environment, policies and markets), knowledge and behavioural outcomes at the household/community/market levels, and impact outcomes, see Figure 1. This is the hypothesised generic impact pathway for IAR4D. Impact pathways for individual SSA CP task forces exhibit minor variations to Figure 1, depending upon the specificities of the problem/opportunity that they address.

The main outcomes at the Innovation Platform (IP) level are increased awareness, increased knowledge drawn from several IP sources, increased access to information, inputs and output markets, and behavioural changes at the individual and system level. These combine to generate innovations directly and at the interfaces of productivity, care for the environment, policies, markets, product development, nutrition and gender, with a potential to demonstrably increase the delivery of benefits to end users. This will, in turn, lead to outcomes at farm household, village community, and market levels. The main outcomes at the household and community levels are as follows:

- increased awareness and knowledge;
- behavioural outcomes (such as adoption of relevant innovations, more effective supply of inputs to satisfy demand, increased and better expressed demand for inputs, and increased volume of input sales);
- market outcomes (increased and more effective supply of outputs, increased demand by consumers); and
- efficiency outcomes (increased yields, technical and allocative efficiency and profit).

These outcomes lead to impacts in the form of welfare and equity outcomes (such as increased incomes, poverty reduction, improved health and nutrition, and equity) and environmental outcomes (for example, imputed soil fertility and erosion). It is hypothesized that evidence provided by the SSA CP's research comparing the benefits of IAR4D against conventional ARD approaches will determine whether communities and other organizations more directly involved in development will seek to adopt and use the IAR4D approach and further scale it out to meet their needs. The outcomes and range of IAR4D's impact are influenced by several conditioning factors (see Figure 1). These factors complicate the attribution of changes in impact indicators to IAR4D alone. Factors exogenous at the household level but endogenous at the community level include infrastructure (public and privately supplied), institutions (governance and market structures), policies (macroeconomic, sectoral, pricing, social), technologies and information. These factors are well anticipated in the formation of Innovation Platforms as fora bringing together players that can potentially make necessary changes that may lead to the removal of obstacles against use of research results. Factors exogenous at the community level include agro-climatic conditions and external market conditions (world prices and access to foreign markets).

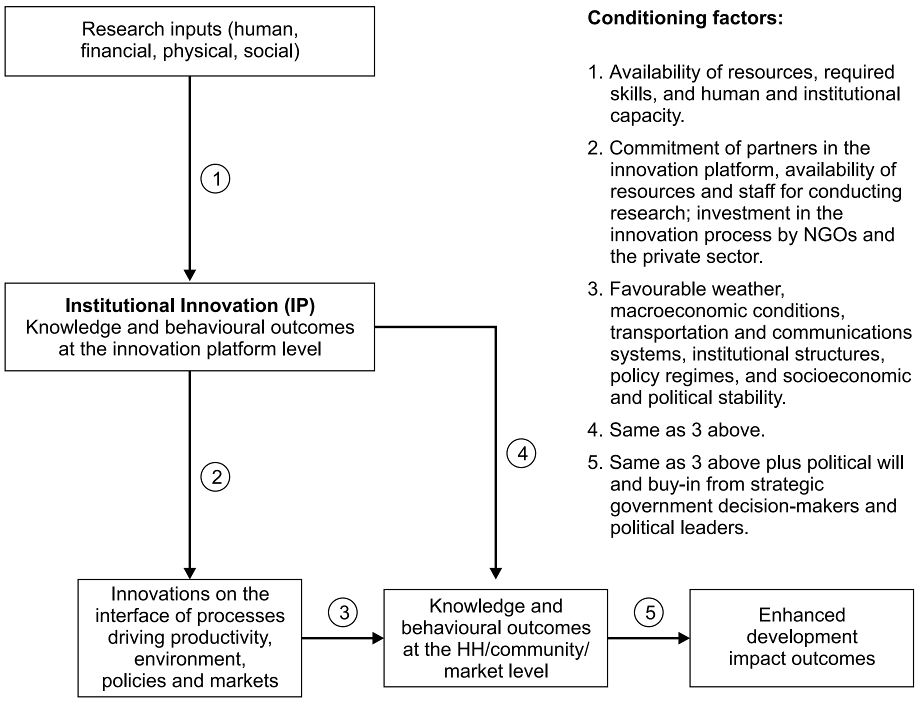



***Figure 1. IAR4D impact pathway.***

(Source: FARA 2009, Figure 3, p. 11)

## Research questions and hypotheses

The SSA CP tested three hypotheses that flow from the three research questions, as shown in Table 1.

**Table 1. Research questions and hypotheses.**

| Research Question | Hypothesis |
|---|---|
| 1. Does the IAR4D concept work and can it generate International Public Goods (IPGs) and Regional Public Goods (RPGs) to end users? | H1: If an innovation platform is created and is functional with the 5 components characterizing IAR4D, then it will lead to increased interactions among partners in the IP, compared to where there is no IP, and increased interactions among farm households in communities and better developmental outcomes where IAR4D is in operation, compared to communities where IAR4D is not in operation. |
| 2. Does the IAR4D framework deliver more benefits to end users than conventional approaches (assuming conventional research, development and extension approaches have access to the same resources)? | H2: IAR4D delivers more benefits to end users and communities compared to conventional approaches (if the conventional ARD approaches have access to the same resources). |
| 3. How sustainable and usable is the IAR4D approach outside its test environment, that is, concerning its scaling out for broader impact? | H3: If IAR4D works in the different PLS contexts, then it can be extrapolated outside the test environments. |

## Analytical Approach

The main aim of this report was to evaluate the impact of Integrated Agricultural Research for Development (IAR4D) on the key outcomes of the implementation of the SSA CP. These outcomes include, among others, poverty reduction and food security. The SSA CP's IAR4D is being implemented through the Innovation Platform (IP) systems in three Pilot Learning Sites (PLS) of three sub-regions of sub-Saharan Africa, namely, Kano-Katsina-Maradi (KKM PLS) in West Africa; Lake Kivu (LK PLS) in East/Central Africa; and Zimbabwe-Malawi-Mozambique (ZMM PLS) in Southern Africa. Each of the three PLSs is made up of three Task Forces (the task forces are described in the PLS reports). In all, the programme thus has nine task forces, which are characterized on the basis of agroecogical parameters, market opportunities and other features. Each of the Task Forces is made up of 4 IPs. The programme is, therefore, made up of 36 IPs. For each of the IPs where the SSA CP's IAR4D is intervening (the treated site), there are two control sites, namely the conventional ARD and the clean sites. In other words, the IPs are the treated sites and the ARD and the clean sites are the non-treated sites. The IPs are treated with the IAR4D, where existing and/or new technologies are being promoted. If the technologies were randomly assigned to farmers, we could assess the impact of their adoption on households' food security and poverty levels by comparing the average outcomes of the treated and the non-treated households. In such a case, the average treatment effect (ATE) can be computed as follows:

$$ATE = E(Y_i \mid \mathrm{D} = 1) - E(Y_0 \mid D = 1) \qquad (1)$$

This is based on the assumption that the outcome levels of the treated before the intervention of the IAR4D $E(Y_0 \mid D = 1)$ can reasonably be approximated by the outcome level of the non-treated during data collection $E(Y_0 \mid D = 0)$. Otherwise, estimation of ATE using the above equation is not possible, since we do not observe $E(Y_1 \mid D = 1)$ though we do observe $E(Y_0 \mid D = 1)$ and $E(Y_0 \mid D = 0)$. However, technologies are rarely randomly assigned. Instead, technology adoption usually occurs through the self-selection of farmers or, sometimes, through programme placement. In the presence of self-selection or programme placement, the above procedure may result in a biased estimation of the impacts of improved technologies, since the treated group (i.e., the IAR4D site—IP farmers) are less likely to be statistically equivalent to the comparison group (i.e., the ARD and clean site farmers) in a non-randomized setting. The propensity score matching (PSM) method, which was developed by Rosenbaum and Rubin (1983), has been extensively used in economics since the 1990s to solve that problem. Rosenbaum and Rubin (1983) defined 'propensity score' as the conditional probability of receiving a treatment given pre-treatment characteristics:

$$P(X) \equiv \Pr\{D = 1 \mid X\} = E\{D \mid X\} \qquad (2)$$

where $\mathrm{D} = \{0, 1\}$ is the indicator of exposure to treatment and $\mathrm{X}$ is the multidimensional vector of pre-treatment characteristics.

The PSM method is a systematic procedure of estimating counterfactuals for the unobserved values $\mathrm{E}\ (\mathrm{Y1|D{=}0})$ and $\mathrm{E}\ (\mathrm{Y0|D{=}1})$ to compute the impact estimates with no (or negligible) bias. The validity of the outputs of the PSM method depends on the satisfaction of two basic assumptions, namely: the Conditional Independence Assumption (CIA) and the Common Support Condition (CSC) (Becker and Ichino 2002). CIA (also known as Unconfoundedness Assumption) states that the potential outcomes are independent of the treatment status, *given* $X$. Or, in other words, after controlling for $X$, the treatment assignment is "as good as random". The CIA is crucial for correctly identifying the impact of the programme, since it ensures that, although treated and nontreated groups differ, these differences may be accounted for in order to reduce the selection bias. This allows the nontreated units to be used to construct a counterfactual for the treatment group. The CSC entails the existence of sufficient overlap in the characteristics of the treated and nontreated units to find adequate matches (or a *common support*). When these two assumptions are satisfied, the treatment assignment is said to be strongly ignorable.

**Estimating Propensity Scores and Assessing Match Quality:** We used the probit model to estimate propensity scores. Selected socio-economic and demographic selected variables were included in the model. Because the matching procedure conditions on the propensity score but does not condition on individual covariates, one must check that the distribution of variables are 'balanced' across the treated and non-treated groups. Rosenbaum and Rubin (1985) recommend that standardized bias (SB) and *t*-test for differences be used to check matching quality. If the covariates $X$ are randomly distributed across the treated and non-treated groups, the value of the associated pseudo-$R^2$ should be fairly low and the likelihood ratio should also be insignificant. A bootstrapping method was used to compute the standard error for the estimate of the IAR4D impact.

**Choosing a Matching Algorithm:** Three commonly used matching algorithms, namely nearest neighbour matching, radius matching and kernel-based matching, were employed to assess the impact of IAR4D on households' income. The nearest neighbour matching (NNM) method matches each farmer from the treated group with the farmer from the non-treated group having the closest propensity score. The matching can be done with or without replacement of observations. NNM faces the risk of bad matches if the closest neighbour is far away. This risk can be reduced by using a radius matching (RM) method, which imposes a maximum tolerance on the difference in propensity scores. However, some treated units may not be matched if the dimension of the neighbourhood (i.e., the radius) is too small to contain control units. The kernel-based matching (KM) method uses a weighted average of all farmers in the adopter group to construct a counterfactual. The major advantage of the KM method is that it produces ATT estimates with lower variance since it utilizes greater information; its limitation is that some of the observations used may be poor matches.

## Evaluation Design

In order to test the three hypotheses already mentioned in a statistically robust fashion and empirically determine whether IAR4D works and whether it delivers more benefits than conventional approaches, a multiple treatments experimental design was used. This design compared household and community level outcomes under (i) IAR4D, (ii) the conventional approach, and (iii) no intervention. In other words, the SSA CP experiment comprised three treatments carried out in three blocks (the PLS) and nine repetitions (three per block—the task forces).

Following White and Chalak (2006), the set of *counterfactuals* was taken to be the set of all possible states of the world, with outcomes taking different values under different possible states of the world. An *intervention* was also seen as the move from one possible state to another. So there are as many counterfactuals as there are possible states of the world.[1] However, under the SSA CP we are limiting ourselves to comparing outcomes under IAR4D and under only two other possible states, namely, the conventional approach and under non-intervention. So, our set of counterfactuals is limited to the set $\{\omega_0, \omega_1, \omega_2\}$ where $\omega_0$ is the non-intervention state consisting of having neither IAR4D nor the conventional approach in operation, $\omega_1$ the state consisting of having the conventional approach in operation, and $\omega_2$ is the state consisting of having IAR4D in operation[2].

The effectiveness and impact of IAR4D were assessed throughout the impact pathway all the way to the farmer level. The hypothesis about whether IAR4D works was tested by comparing the values of relevant knowledge, behavioural, efficiency, welfare, equity and environmental outcomes under $\omega_2$ and under $\omega_0$. Similarly, the hypothesis about whether IAR4D delivers more benefits than the conventional approach was tested by comparing the values of relevant knowledge, behavioural, efficiency, welfare, equity and environmental outcomes under $\omega_2$ and under $\omega_1$. The "with" and "without" IAR4D comparison was made by comparing the values of the same outcomes as above under $\omega_2$ and under the composite possible state "$\omega_0$ or $\omega_1$".

## Characterisation of Treatment and Counterfactual Sites

Innovation Platforms were evaluated at the district/local government areas/communes levels because it is conceptualized that the innovation process is best organized through geographically decentralized sites. The geographical area of influence of the IP is conceptualized to be mostly within district/local government areas/commune jurisdictional boundaries, because of the clustering of activities and interactions among government administrative units, public research and extension organizations, farmers, farmers' organizations, NGOs, agricultural input suppliers and output marketing firms, credit and finance organizations, and service providers.

- The treatment communities consist of organizations and farm households in areas where IAR4D was practised.
- The non-treatment communities consist of similar organizations and households in other sites.

1. But among all the possible states of the world, only one gets realized (the factual) in any given situation, all the other are counterfactuals.
2. Only one of the three possible states gets realized in any given site. The realized state will then be the factual and the unrealized ones the counterfactuals.

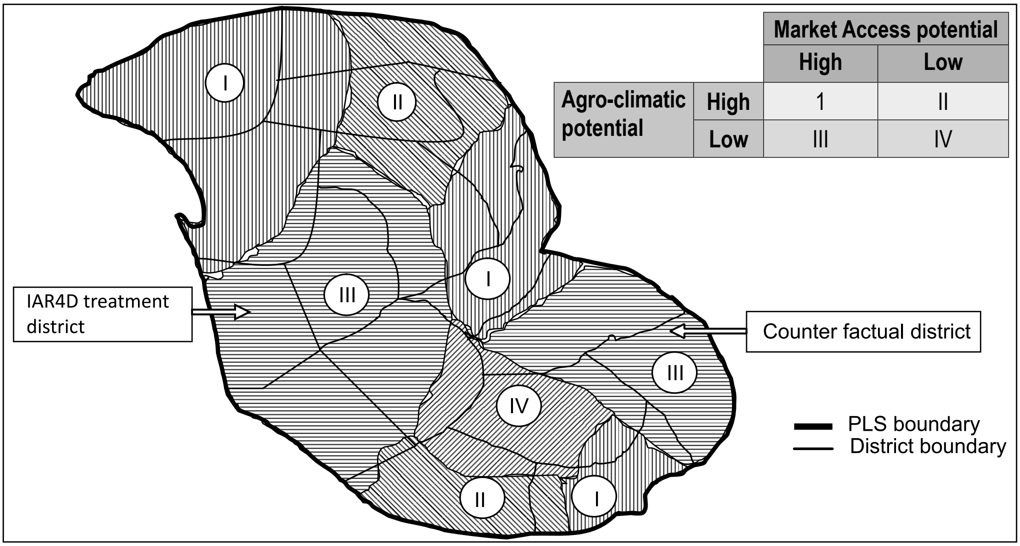



***Figure 2. Illustration of stratification of a PLS by four development domains.***

The PLS was zoned into development domains—areas with comparable development potential. The development domains used by the SSA CP are based on two factors that usually have the largest influence on agriculturally driven development, namely agro-climatic potential and access to markets. The development domains combined with population data were used to target areas most likely to provide the highest returns on the SSA CP's investment. They also provided a basis for stratifying the PLS, in order to capture its variation and to delineate similar domains from which comparable sites were selected.

Research sites (districts/communes/local government areas) were allocated to IAR4D and non-IAR4D treatments through stratified random sampling. The strata within which the randomization was carried out are four development domains, delineating the combination of market access potential and agro-climatic potential. Each IAR4D treatment site (district/commune/local government area) has a corresponding counterfactual site randomly selected from the same stratum as the IAR4D site (for example, see Figure 2). The task forces were spread over IAR4D treatment sites across various strata, in order to investigate the performance of the approach across a wide range of conditions.

Each task force established 4 Innovation Platforms in 4 separate districts/communes/local government areas. For each IP/IAR4D site, a matching comparator site was selected. Thus, each task force worked in 8 districts/communes/local government areas.

Within IAR4D and non-IAR4D sites, focal villages were also selected randomly. The focal villages were screened prior to implementation of IAR4D to establish whether or not they have had conventional ARD or IAR4D-type of projects in the past 2-5 years. Villages were classified into 2 types: (a) "clean" villages that have neither had IAR4D nor conventional projects in the last

3. Whereas "clean" villages are defined as those that have not had any intervention /initiative (conventional or IAR4D-like) over the last 2-5 years, the categorization of any village as "clean" will depend on the local context.

2-5 years[3]; and (b) conventional ARD villages that have had projects identifying, promoting and disseminating technologies in the past 2-5 years.

IAR4D was introduced in "clean" villages within the IAR4D sites. The SSA CP's hypotheses were tested by determining whether outcomes differ among households in the IAR4D, "clean" and conventional ARD villages.

## Census of District/Local Government Area/Commune Sites and Characteristics

A list of districts within the PLS and their characteristics, including their representation of the four development domains, was used to identify clusters for targeting under the research and classification of villages into "clean" and "non-clean" status. Some of this information was collected from national statistical offices and used for stratification and multi-stage sampling of villages. Field visits were made to sampled villages to collect information on their "clean" or "non-clean" status, using key focus group discussions and informant interviews.

An alternative method entails using the geographical information system (GIS) and quota sampling to select district/local government area/commune sites for assignment to treatment. GIS tools were used to randomly pick 5 points, followed by field visits to check the "clean" or "non-clean" status of the quadrant in which they fell. If the quadrant was not "clean", then it was discarded and another one sampled until the quota was met. The advantage of this method is that it does not require a prior census.

## Sampling Method

Multi-stage stratified random sampling was carried out within the selected districts (IAR4D and counterfactual) to select the villages where the treatments were applied, that is villages where IAR4D was introduced, village communities where conventional approaches were in operation, and villages where no interventions had been carried out over the last 2-5 years.

The Miguel and Kremer (2004) method of randomizing treatments across schools (districts and village communities) and not individual farm households was used, because it captures benefits from spillovers and externalities that would be underestimated if the treatment is randomized only at the individual level. All districts/local government areas/communes within the PLS were first listed and grouped according to their representation of the four development domains. Depending on the context and its specific requirements, each task force was defined, the strata within which it randomly selected the four districts served as its IAR4D treatment sites, that is, where IAR4D was introduced. Within the IAR4D sites, a census of the village communities was conducted to develop a village sampling frame and stratify the villages into clean and non-clean villages. At least 5 focal villages per IAR4D site were randomly selected from only clean villages. These villages became the theatres for action research, aimed at developing innovations on the interface between productivity, care of the environment, policies and markets. Within the focal IAR4D village communities, at least 10 households per village were randomly selected for monitoring and evaluation.

Four counterfactual districts/local government areas/communes that were similar to the IAR4D sites (for example, sharing the same development domain) were assigned to conventional and non-IAR4D-non-conventional ("clean" village) treatments. As for IAR4D sites, a village census was carried out and villages stratified into clean and non-clean. For each counterfactual site matching an IAR4D site, 5 focal villages were randomly selected from clean villages only and assigned to the non-IAR4D-non-conventional treatments. Similarly, 5 focal villages were randomly sampled from non-clean villages and assigned to the conventional approach treatment. At least 10 households per focal village were randomly selected for monitoring and evaluation.

## Sample Sizes at IP, Task Force, PLS and SSA CP Scales

**Sample sizes for task forces.** For each of the three treatments (IAR4D, conventional ARD and no intervention at all), task forces initiated action research in 5 focal villages. Thus each task force (4 sites) worked in 60 villages. Within each village, the task forces monitored 10 households. Over the 60 IAR4D villages, each task force monitored 600 households.

**Sample sizes for PLS and Programme.** At the PLS level, the task force sample sizes were multiplied by a factor of 3, reflecting the three task forces in each PLS. At programme level, the task force sample sizes were multiplied by a factor of 9, reflecting the nine task forces constituting the programme, that is, 36 Innovation Platforms sites and 36 comparator (counterfactual) sites; 540 villages, and 5,400 households consisting of 1,800 IAR4D treatment villages, 1,800 conventional and 1,800 non-IAR4D-non-conventional counterfactual villages.

The sample sizes at the four scales, that is, IP, task force, PLS and programme are summarised in Table 2.

**Table 2. Sample sizes at IP, task force, PLS and programme scales.**

| | No. of IPs | No. of IAR4D villages | No. of conventional ARD villages | No. of non-IAR4D-non-conventional villages | Total number of households |
|---|---|---|---|---|---|
| IP | 1 | 5 | 5 | 5 | 150 |
| Task force | 4 | 20 | 20 | 20 | 600 |
| PLS | 12 | 60 | 60 | 60 | 1800 |
| SSA CP | 36 | 180 | 180 | 180 | 5400 |

Observations and results across task forces were pooled using meta-modeling, to evaluate the site-to-site variation of IAR4D treatment effects. Observations and results were pooled across multiple sites and PLS to evaluate the programme and predict the impact of IAR4D at new sites.

## Sample Selection

The data used in this report were taken from baseline and midline surveys of over 5,400 households across the SSA CP. The survey was conducted by task forces within the framework of the SSA CP, supported by the Forum for Agricultural Research in Africa (FARA) and its donors—including the European Commission (EC), the UK Department for International Development (DFID UK), and the Governments of Italy and Norway.

The sample frame was derived from different districts, selected to represent the basic areas of task forces in the three PLS that made up the SSA CP. In each district, a sample of households was selected by taking a sample of district wards; a random sample of villages within each ward; and a random sample of households in each selected village. Finally, a household was retained in the sample if it belonged to one of the 540 villages selected within the clean, conventional or IP/action sites.

## Baseline Surveys for IP and Community Level Characteristics

Baseline surveys, field observations and focus group discussions were conducted to benchmark pre-treatment characteristics of IPs, site characteristics and baseline levels of outcomes predicted under the IAR4D approach: number, variety and time to develop innovations; knowledge and behavioural outcomes (adoption, input supply, input demand, volume of sales); market outcomes (output supply and consumption demand); productivity outcomes (yields, technical and allocative efficiency, and profit) and impacts (incomes, livelihood assets and equity). Several indicators were used to measure outcomes, which were different with context. The questionnaires were designed for comparison within an IP over time and across IPs. To generate counterfactuals, surveys and field observations were conducted in the comparison sites and villages assigned to conventional and non-IAR4D-non-conventional treatments. Key players in the innovation systems—such as public and private agricultural researchers, extension workers, farmer leaders, traders, dealers, lenders and key informants—were interviewed to characterise innovation systems and establish the baseline levels in the IP sites.

## Baseline Survey for Household and Village Community Characteristics

Baseline surveys, observations and focus group discussions were conducted to collect data on household-level and village-community-level characteristics, and behavioural, efficiency, environmental and welfare outcomes. Surveys were used to track feedback, information diffusion, awareness and knowledge changes, adoption, and market effects of innovations and spillovers, using the Miguel and Kremer (2004) approach and other methods.

## Evaluation Surveys

Follow-up evaluation surveys and qualitative assessment studies were conducted in the third year (2010) to assess the implementation process; document all the intermediate steps of the research-to-impact pathway and conditioning factors; assess participants' subjective reactions to IAR4D; identify subgroups experiencing greater or lesser impact than the sample as a whole; and measure changes in outcomes at the levels of the IP, household, community and market. Follow-up surveys used the same indicators as were used in the baseline surveys to measure outcomes.

## Data Analysis

Assessing the impact of any intervention requires making an inference about the outcome that would have been observed had the programme participants not participated. Following Heckman et al. (1997) and Smith and Todd (2001), let $\mathrm{Y}_1$ be the mean of the outcome conditional on participation, that is, treatment group, and let $\mathrm{Y}_0$ be the outcome conditional on non-participation, that is control group. The impact of participation in the programme is the change in the mean outcome caused by participating in the programme, which is given by

$$\Delta Y = Y_1 - Y_{0,} \quad \text{.........................................................} (1)$$

where $\Delta$ is the notation for the impact for a given household (1)

The fundamental problem of evaluating this individual treatment effect arises because for each household, only one of the potential outcomes, either $\mathrm{Y}_1$ or $\mathrm{Y}_{0,}$ can be observed, but $\mathrm{Y}_1$ and $\mathrm{Y}_0$ can never be observed for the same household simultaneously. This leads to a missing-data problem, which is the heart of the evaluation problem (Smith and Todd 2001). The unobservable component in equation (1), be it $\mathrm{Y}_1$ or $\mathrm{Y}_0$, is called the counterfactual outcome. Measuring impact as the difference in mean outcome between all households involved in the project and those not involved, even when controlling for programme characteristics, may thus give a biased estimate of programme impact. Since there will never be an opportunity to estimate individual treatment effects in (1) directly, one has to concentrate on population averages for the impacts of a treatment.

Two treatment effects are dominantly used in empirical studies. However, the most commonly used evaluation parameter is the so-called average impact of the treatment on the treated (ATT), which focuses explicitly on the effect on those for whom the programme is actually introduced. In a random programme assignment, the expected value of ATT is defined as the difference between expected outcome values with and without treatment for those who actually participated in the treatment (Heckman et al. 1998b), which is given by

$$\Delta Y_{ATT} = ATT\ (\Delta Y|\ X\!: Z = 1) = E(Y_1 - Y_0|,\ Z = 1) = E\ (Y_1|\ Z = 1) - E(\ Y_0|\ Z = 1)\ ...(2)$$

where Z is an indicator variable, indicating whether a household i actually received treatment or not: Zi being equal to 1 if the household is a beneficiary, and 0 otherwise. X denotes a vector of control variables.

Data on programme beneficiaries identify the mean outcome in the treated state E (Y1|X, Z=1). The mean outcome in the untreated E (Y0|X, Z=1) is not observed, and a proper substitute for it has to be chosen in order to estimate ATT.

Various quasi-experimental and non-experimental methods have been used to address the bias problem (Heckman et al. 1998 a). One of the most commonly used quasi-experimental methods is propensity score matching (PSM), which selects project beneficiaries and non-beneficiaries who are as similar as possible in terms of observable characteristics expected to affect project participation as well as outcomes. The difference in outcomes between the two matched groups can be interpreted as the impact of the project on the beneficiaries (Smith and Todd 2001). We used this method to estimate the ATT for impacts of the IAR4D on the key outcomes of the project (that is, poverty/food security, factor productivity, market participation, awareness and adoption, as well as natural resource management).

The PSM method matches project beneficiaries with comparable non-beneficiaries using a propensity score, which is the estimated probability of being included in the project. Only beneficiaries and non-beneficiaries with comparable propensity scores are used to estimate the ATT. Those who do not have comparable propensity scores are dropped from the comparison groups.

Among the advantages of PSM over econometric regression methods is that it compares only comparable observation and does not rely on parametric assumption to identify the impacts of projects. However, PSM is subject to the problem of “selection on unobservables”, meaning that the beneficiary and comparison groups may differ in unobservable characteristics, even though they are matched in terms of observable characteristics (Heckman et al. 1998a). Econometric regression methods devised to address this problem suffer from the problems previously noted. The bias resulting from comparing non-comparable observations can be much larger than the bias resulting from selection on unobservables, although we cannot be certain whether that conclusion holds in general (Heckman et al. 1998a).

In this study, we address the problem of selection on unobservables by combining PSM with the use of the double-difference (DD) estimator. The double-difference estimator compares changes in outcome measures (i.e., change from before to after the project) between project participants and non-participants, rather than simply comparing outcome levels at one point in time.

$$DD = (Y_{p1} - Y_{p0}) - (Y_{np1} - Y_{np0})\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(3)$$

where $Y_{p1}$ = outcome (e.g., income) of beneficiaries after the project started; $Y_{p0}$ = outcome of beneficiaries before the project started; $Y_{np1}$ = outcome of non-beneficiaries after the project started; and $Y_{np0}$ = outcome of non-beneficiaries before the project started.

The advantage of the double-difference estimator is that it nets out the effects of any additive factors (whether observable or unobservable) that have fixed (time-invariant) impacts on the outcome indicator (such as the abilities of the farmers or the inherent quality of natural resources), or that reflect common trends affecting project participants and non-participants equally (such as changes in prices or weather; see Ravallion 2005).

Thus, for example, if project participants and non-participants are different in their asset endowments (mostly observable) or in their abilities (mostly unobservable), and if those differences have an additive and fixed effect on outcomes during the period studied, such differences will have no confounding effect on the estimated ATT.

In principle, the double-difference approach can be used to assess project impacts without using PSM, and it will produce unbiased estimates of impacts as long as these assumptions hold. However, if the project has differential impacts on people with different levels of wealth or observable characteristics, the simple double-difference estimator will produce biased estimates if participant and non-participant households differ in those characteristics (Ravallion 2005). By combining PSM with the double difference estimator, controls for differences in pre-project observable characteristics can be established. A bias could still result from the heterogeneous or time-variant impacts of the unobservable differences between participants and non-participants. For example, communities and households that had participated in ARD may have different responses to IAR4D than those in the clean environment, because of the cumulative effects of social capital developed under the ARD, favourable or adverse

**Table 3. Variables used to compute propensity scores and their expected signs.**

| Variable | Expected impact on participation in IAR4D | Why? | Expected sign on income and wealth | Why? |
|---|---|---|---|---|
| Gender of Respondent (Male=1; Female=0) | – | IAR4D is gender friendly | – | Women are usually poorer than men |
| Household Size | + | Larger families could be associated with poverty or other vulnerabilities that makes participation in IAR4D worthwhile | – | The larger the family, the poorer it is |
| Age of respondent | +/– | IAR4D supports both the young and old | + | Older respondents likely to be better off because of accumulation of wealth and experience over the life cycle |
| Level of Education of respondent(years of formal education) | + | Some project requirements need a certain level of education | + | Education increases income opportunities, such as on-farm activities |
| Area of farmland cultivated (ha) | +/– | IAR4D concept encourages more area of land to be cultivated | + | More area of land enables households to earn more income and more productive assets |
| Agro-ecological Zone | +/– | The technologies promoted by IAR4D in each agro-ecology motivate participation | - | Some zones closer to urban centers have more potential of membership than the remote ones |
| Distance to nearest all weather road | + | Closeness to urban center encourages participation since products are easily marketed | + | Access to improved road increases income opportunities and reduces transaction costs |
| Value of productive asset | + | Same as for land area | + | Same as for land area |

Source: Data Analysis 2012

experiences under ARD, or other factors. Such shortcomings are unfortunately inherent in all non-experimental methods of impact assessment (Duflo et.al. 2006). Although no solution to these potential problems is perfect, we believe the method we have used addressed these issues as well as possible in this case.

The standard errors estimated by the double-difference method may be inconsistent because of serial correlation or other causes of a lack of independence among the errors.

In ordinary regression models, serial correlation can result from unobserved fixed effects, but by taking first differences, the double-difference method eliminates that source of serial correlation. However, serial correlation still may be a problem if more than two years of panel data are used (Duflo et al. 2004). In this study, because we used only two periods, before and after the project, we do not have the concern about serial correlation among multiple periods. Another reason for the possible non-independence of the errors is clustering of the sample.

The propensity scores were computed using binary logit regression models. We estimated three probit models for three comparisons: (1) IAR4D beneficiaries compared with all non-beneficiaries; (2) IAR4D beneficiaries compared with conventional beneficiaries, and (3) IAR4D beneficiaries compared with non-beneficiaries in clean communities. The dependent variable in each model is a binary variable, indicating whether the household was a beneficiary of the IAR4D project.

The explanatory variables used in computing the propensity scores are those expected to jointly determine the probability to participate in the project and the outcome. We focused on the determinants of income and productive assets when selecting the independent variables for computing the propensity score matching.

The independent variables used in the regression are summarized in Table 3.

# Chapter 3
# Results and Discussion

The discussion of results here is divided into two parts: (1) the impact of IAR4D on household income; and (2) the estimation results from the probity scores.

## Impact of IAR4D on Household Income

The 2008 average income for treated (clean before intervention), conventional and the clean sites were US$ 588.49, $ 764.06 and $672.80, respectively. At midline, the average incomes were estimated to be US$ 1358.57; $824.01 and $752.88, respectively (Table 4). The average treatment effect on the treated (ATT) was computed based on two alternative matching methods. The outcome variable is household per capita income per year measured in US Dollars. The z-statistics were based on bootstrapped standard errors, with 50 replications, which were used to verify whether the observed effect was significant or not.

The results show that the average income of the treated (IAR4D farmers) sample due to participation in the IP activities based on the PSM (ATT) was US$ 1362.72 in the case of both the Kernel and nearest neighbour matching estimates ($p < 5\%$). A comparative analysis shows that the IP farmers are better than the farmers in the two counterfactuals of conventional and clean sites.

## Estimation Results of Propensity Scores

The importance of estimation of propensity scores is twofold: first, to estimate the ATT and, second, to obtain matched treated and non-treated observations. The results of the probit models are reported in Table 5.

Table 4. Impact of IAR4D on household income across types of respondents.

| | Net Real Household Income (US $) | | ATT | % change due to participation in IAR4D |
|---|---|---|---|---|
| | Before IAR4D | After participation in IAR4D | | |
| IAR4D (n=1442) | 588.49 (42.91) | 1358.37 (84.82) | 1362.72** (614.11) | 231.56 |
| Conv (n=1563) | 664.06 (49.90) | 824.01 (95.48) | -1013.91 (470.22) | |
| Clean (n=1465) | 672.80 (45.70) | 752.88 (544.71) | -488.54 (452.21) | |
| **Gender (female only)** | | | | |
| IAR4D n=582 | 350.51 (77.24) | 620.38 (117.26) | 1141.10*** (465.29) | 325.54 |
| Conv n=647 | 465.86 (203.21) | 534.49 (273.45) | -682.74* (402.57) | |
| Clean n=647 | 487.43 (111.74) | 588.39 (148.19) | -617.59 (522.95) | |
| **Food Security** | | | | |
| IAR4D n=813 | 446.00 (49.08) | 1103.14 (76.60) | 1444.86*** (536.71) | 323.96 |
| Conv n=839 | 522.06 (44.90) | 360.61 (105.96) | -981.94 (629.55) | |
| Clean n=811 | 411.16 (40.82) | 700.24 (137.32) | -759.82 (646.35) | |
| **Research** | | | | |
| IAR4D n=97 | 387.51 (100.00) | 707.04 (112.55) | 1510.42 (658.46) | 389.78 |
| Conv n=101 | 229.07 (195.41) | 328.92 (199.90) | -995.71 (613.24) | |
| Clean n=99 | 274.49 (58.51) | 521.27 (175.86) | -665.91 (529.22) | |
| **Youths** | | | | |
| IAR4D n=893 | 417.76 (43.22) | 873.62 (73.96) | 1541.08* (1002.93) | 368.89 |
| Conv n=947 | 544.56 (53.52) | 591.05 (100.70) | -958.18 (732.54) | |
| Clean n=480 | 120.73 (109.93) | 345.43 (234.47) | -480.55 (680.94) | |
| **Wealth Distribution** | | | | |
| **Tercile1 (poorest)** | | | | |
| IAR4D n=497 | 744.09 (33.43) | 3021.70 (212.52) | 39.82* (30.71) | 5.35 |
| Conv n=424 | 949.06 (39.42) | 1892.28 (248.47) | -42.51 (39.80) | |

Continued...

*Table 4 Continued...*

| | | | | |
|---|---|---|---|---|
| Clean<br>n=484 | 893.43<br>(34.74) | 2995.21<br>(162.70) | -1.56<br>(39.33) | |
| **Tercile2** | | | | |
| IAR4D<br>n=284 | 577.59<br>(196.91) | 658.10<br>(130.90) | 63.61**<br>(34.79) | |
| Conv<br>n=395 | 638.42<br>(172.93) | 614.21<br>(225.33) | -17.73<br>(51.68) | 4.97 |
| Clean<br>n=256 | 722.57<br>(230.73) | 576.36<br>(384.70) | -42.78<br>(40.62) | |
| **Tercile3** | | | | |
| IAR4D<br>n=661 | 175.42<br>(20.10) | 408.61<br>(27.42) | 121.81<br>(124.63) | 69.44 |
| Conv<br>n=744 | 194.42<br>(15.08) | 289.44<br>(41.59) | -104.52<br>(144.36) | |
| Clean<br>n=725 | 154.84<br>(13.78) | 250.93<br>(27.28) | -16.38<br>(107.48) | |

ATT = (Yp1-Yp0)-(Ynp1-Ynp0). "Before project" is the situation before the IAR4D in 2008, while "After project" is two years after the project started in 2010.

"ATT" and the corresponding "%" refers to the change in measured household income resulting from participation in the Innovation Platform (IP) of the IAR4D. % net change due to participation at the platform = (ATT/Yp0)*100.

Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

The results of the probit regression (Table 5) show that the participants in the IAR4D would most likely be young married farmers with small family size, with the need to consciously encourage farmers from Malawi, Mozambique, and Zimbabwe to participate. However, participants in the conventional module are mostly farmers with some productive assets (here farmers from Zimbabwe need to be consciously encouraged to participate), while those in the clean sites are female farmers without productive assets (and here the farmers from Mozambique need to be encouraged to participate). This result suggests that the IAR4D intervention focused on married youths, who are the more vulnerable groups in sub-Saharan Africa.

These probit model results were used to compute the propensity scores that were used in the PSM estimation of ATT. Several methods are possible for selecting matching observations (Smith and Todd 2001). We used both the kernel matching method (using the normal density kernel), which uses a weighted average of "neighbours" (within a given range in terms of the propensity score) of a particular observation to compute matching observations, as well as the nearest-neighbour method; using a weighted average improves the efficiency of the estimator (Smith and Todd 2001). Observations outside the common range of propensity for both groups (i.e., lacking "common support") were dropped from the analysis. This requirement of common support eliminated more than half of the total number of observations, indicating that many of the observations from various strata were not comparable.

**Table 5. Probit regression of IAR4D participation (matched observations).**

| Explanatory variables | Treated (IAR4D) | | Control (Conventional) | | Control (Clean) | |
|---|---|---|---|---|---|---|
| | Coefficient | Standard Error | Coefficient | Standard Error | Coefficient | Standard Error |
| Gender (1=male; 0=female) | -0.189 | 0.215 | -0.257 | 0.215 | 0.393 | 0.232** |
| Age of respondent (yrs) | -0.619 | 0.272** | 0.188 | 0.290 | 0.390 | 0.292 |
| Marital Status (categorical) | -0.093 | 0.084 | 0.099 | 0.090 | 0.101 | 0.091 |
| Educational Status | -0.174 | 0.084** | 0.027 | 0.085 | 0.128 | 0.086 |
| Household size | -0.186 | 0.112*** | 0.177 | 0.117 | -0.019 | 0.117 |
| Farm Exp (yrs) | 0.018 | 0.124 | -0.012 | 0.132 | -0.031 | 0.133 |
| Assets (productive) | 0.029 | 0.036 | 0.060 | 0.037* | -0.086 | 0.038** |
| Ngs | -0.078 | 0.303 | -0.162 | 0.313 | 0.213 | 0.311 |
| Sudan | -0.016 | 0.291 | 0.049 | 0.300 | -0.025 | 0.300 |
| Sahel | -0.480 | 0.306 | 0.292 | 0.311 | 0.150 | 0.312 |
| Malawi | -1.761 | 0.356* | | | | |
| Mozambique | -1.393 | 0.367* | -0.259 | 0.359 | -0.761 | 0.367** |
| Zimbabwe | -0.660 | 0.393*** | 0.778 | 0.389** | -0.092 | 0.392 |
| Constant | 2.982 | 0.956 | -2.668 | 1.012 | -2.468 | 1.022 |
| Sample size (n) | 1587 | | 1587 | | 1587 | |
| Pseudo $R^2$ | 0.054 | | 0.019 | | 0.018 | |
| Prob > $\chi^2$ | 0.000 | | 0.001 | | 0.002 | |
| Log likelihood | -868.56 | | -801.16 | | -789.47 | |

Source: Data Analysis (2012)

Further testing of the comparability of the selected groups was done using a "balancing test" (Dehejia and Wahba 2002), which tested for statistically significant differences in the means of the explanatory variables used in the probit models between the matched groups of the IAR4D participants and non-participants. In all cases, that test (balancing test) showed statistically insignificant differences in observable characteristics between the matched groups (but not between the unmatched samples), supporting the contention that the PSM ensures the comparability of the comparison groups (at least in terms of observable characteristics).

We used bootstrapping to compute the standard errors of the estimated ATT, generating robust standard errors, because the matching procedure matched control households to treatment households "with replacement" (Abadie and Imbens 2006).

The experimental design of the project is such that it would allow an examination of spillover effect of the IAR4D, by comparing the changes in income of the participants with those of non-participants living within and outside the communities with the project. The homogenous results suggest that non-participants may have benefited from spillover of the project. For example, non-participants used the innovations and research knowledge made available to the participants, and benefited from policies and infrastructural facilities provided through the intervention of the Innovation Platform. In addition, some services made available to participants could also be available to non-participants; for instance, the storage facilities, the shredding machine, as well as employment, could be made available to non-participants.

It is likely that the impact of the project on incomes will be larger than currently captured because of lagged effects of investments on productive assets, infrastructure and other project investments. The results in Table 4 show the homogenous impact of the IAR4D on the participants' income from data collected after two years. The result shows that participation in IAR4D had a positive and significant impact on the beneficiaries at the 5 % level. The quantum of the impact made the beneficiaries about 230 % better than the baseline condition, while the counterfactual situations (both conventional and clean) were neither better nor statistically significant. The ATT yielded a value of US$ 1362.72, which translated to a daily income of US$ 3.73 per participant. This figure is more than three times the World Bank poverty line of $1.00 per day; in other words, participation in IAR4D enabled participants to cross over the poverty line comfortably. Further analysis showed that with an average household size of 8, the IAR4D must have improved the income of 11,536 people in the two years of its operation.

The effect of the IAR4D varied across the major agro-ecological zones of the PLS. However, the PLS analyses had examined the agro-ecological differences of the concept. Rather, we estimated the hetero-analyses by considering cross-cutting issues, such as gender and food security, among others.

Given the widely acknowledged fact that women and girls are more vulnerable to poverty, a hetero-analysis by gender was carried out and the result in Table 4 shows that the ATT of women was US$ 1141.10, which was significant at 1% and 326% better than at the baseline level. This translates to a daily income of US$ 3.13 for about 600 women participants. In essence, women participants in the IAR4D were able to come out of the poverty trap by their participation, whereas those in both the conventional and clean sites were worse off than at baseline.

In terms of food security, the results show that participants were able to cross over the food insecurity debacle by making an income of about US$ 3.96 daily, which was 324% better than at the baseline level. In terms of number of people, about 6500 became food-secure due to their participation in the IAR4D.

The IAR4D concept was youth sensitive and made participants about 370% better than at the baseline level. The participants had a daily income of about US$ 4.22, thus crossing comfortably the poverty line, while those in the two counterfactuals were worse off. Altogether, the income of about 7144 youths was improved as a result of their participation in the IAR4D.

In terms of wealth distribution, the IAR4D had a significant impact on the poorest tercile of the population, making them 5% better than at the baseline level, whereas the impact on the other two tercii were not statistically significant. In other words, the IAR4D was focused on the poorest segment of the target population, that is, it was pro-poor.

**Table 6. Impact of IAR4D on income distribution.**

| Treatment Type | Gini Coefficient at Baseline | Gini Coefficient at Midline | % Gini Coefficient Change |
|---|---|---|---|
| All respondents | 0.815 | 0.786 | -2.90 |
| IAR4D beneficiaries | 0.798 | 0.737 | -6.10 |
| Conventional | 0.799 | 0.749 | -5.00 |
| Clean | 0.783 | 0.803 | +2.00 |

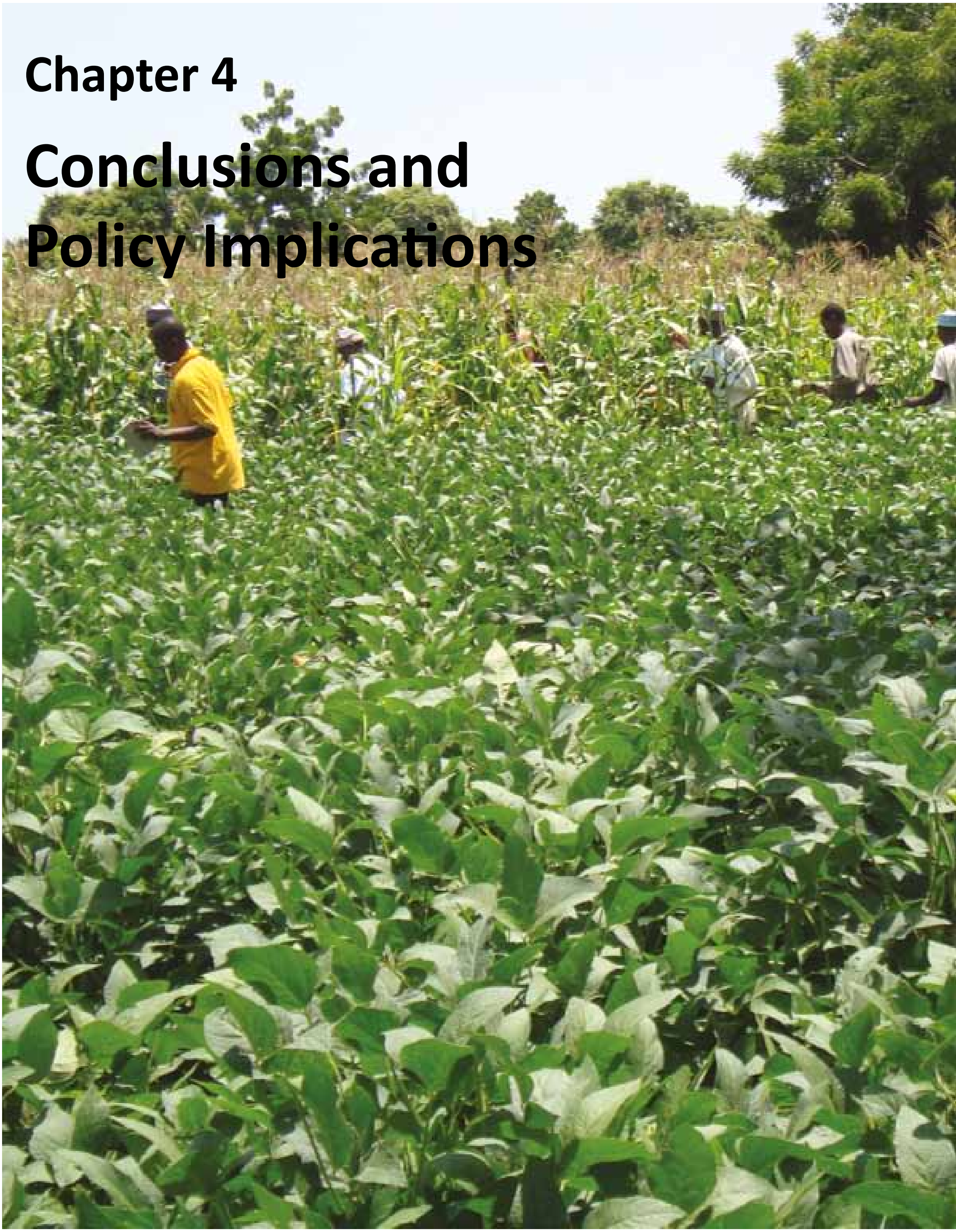
Source: Data Analysis 2012

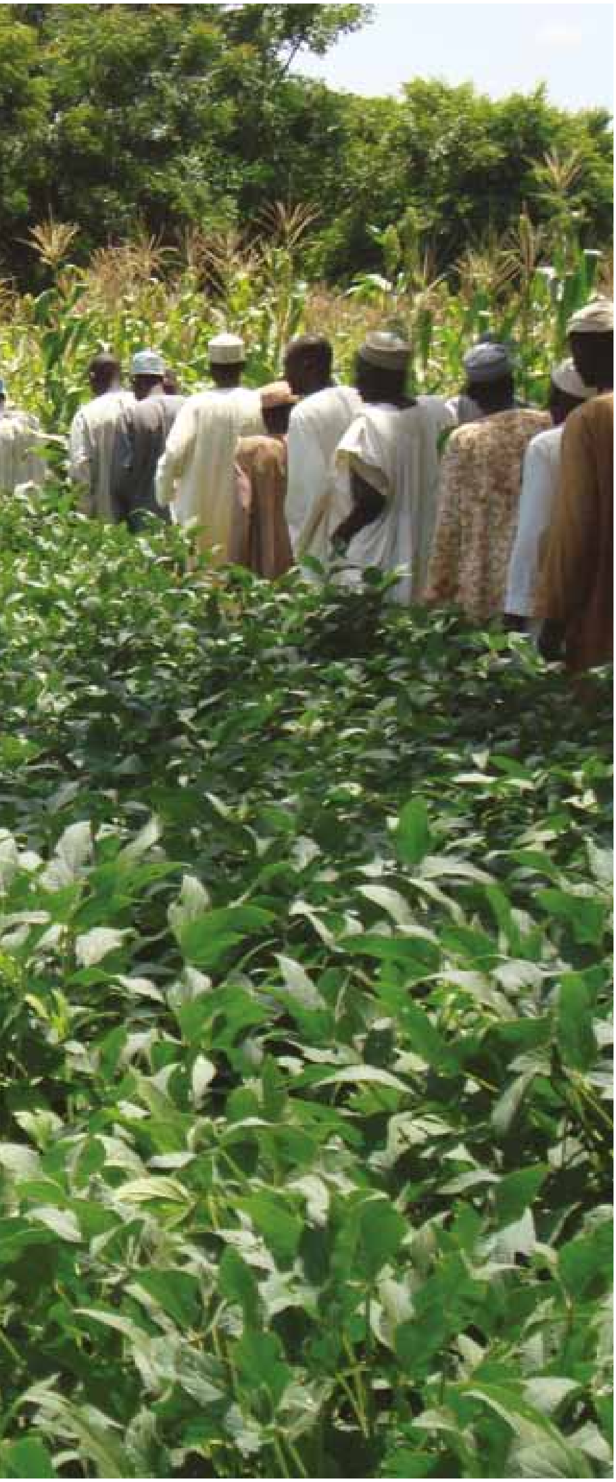
The IAR4D targets the poor and vulnerable groups, such as women, youth, and the elderly. This action is expected to reduce income inequality. The impact of this targeting was examined by considering the change in income inequality over the two years of the project. We computed the Gini coefficient of the income of the respondents for this objective (results in Table 6). Indeed, the results show that the Gini Coeficient of the beneficiaries decreased by about 6 %, suggesting that the project contributed to reduction of income inequality. Income inequality was reduced in the whole project area, as shown by the value of the Gini coefficient being 2.90 percent; however, in the clean zone, there was an increase in income inequality by 2%.

The largest decrease in income inequality is among the IAR4D beneficiaries, showing a figure of about 6.1%, relative to a figure of 5% for the conventional sites. This is consistent with the results that the income of the poorest increased more significantly than that of the middle and upper terciles.

# Chapter 4

# Conclusions and Policy Implications





The proof of concept exercise set out with three questions, aimed at establishing the IAR4D not only as a concept but as a viable alternative to the traditional R&D (conventional approach) which will take Africa's agriculture to the desired level, where the research outputs will be of benefit to both the remote and immediate environment, as well as improve the livelihood of rural farmers in Africa.

We report on each of the questions below, along with the responses our study and analysis have provided us.

## Does the IAR4D work as a concept?

The answer to this question is in the homogenous result of the impact analysis. The results from these studies show that IAR4D as a concept, using Innovation Platforms as the framework, works in delivering development impact. IAR4D works and impacts positively on the lives of the beneficiaries to the tune of US $1362.72 per participant, or US $3.73 per day. This amount lifted about 11,600 people, as direct beneficiaries in the PLS, well above the poverty level in the villages studied.

## Does the IAR4D deliver more benefits than the conventional R&D methods?

With the use of matching methods, as well as the PSM and double difference approach, we can safely conclude from the results that the IAR4D delivers more benefits than the conventional R&D method. The results, while showing the positive impact for the IAR4D, revealed that under the same conditions, the conventional and the clean did not impact consistently positively on the non-beneficiaries.

The aforementioned analyses also showed that the IAR4D impacts on women's income, research participation and food security. In fact, the results showed that about 6500 people were able to cross the food insecurity line, and the income of about 7144 youths was improved as a result of participation in the IAR4D. These results are consistently robust and reliable.

## Can the IAR4D be scaled up and out beyond the current area of operation?

The results of the ex-ante analysis, in line with the impact assessment analysis, suggest that the concept can be successfully scaled up and out, with potentially multiple positive impact on the beneficiaries. Indeed, reports of the success story of the concept abound as to how eager the neighbouring communities are to key into the concept so as to better their lives.

The IAR4D concept had been on the ground for about two years in the three PLS, during which time the project realized significant positive impacts on household income, food security, gender, and research participation. Using propensity score and double-difference methods to control for project placement and self-selection biases, we found that IAR4D increased participants' income, improved household assets, and encouraged participation in research as well as adoption of research outputs.

Household incomes improved substantially more for the IAR4D participants than for non-beneficiaries in conventional and clean sites, with an average increase in real incomes resulting from participation of about 232%, which is not only better than the conventional and clean sites but well above the achievement of similar projects in the continent. For instance, the World Bank sponsored Fadama II project in Nigeria, which won the Banks' Regional Excellence Award, had an income impact rate of about 60%, a feat achieved in six years of operation.

The results of the potential economic surplus model show that the Sudan Savanna region stands to gain an estimated US$ 12 million per year from adoption of the IAR4D approach for maize production. From these benefits, present producer surplus was about US$ 306 million (about 60%)—equivalent to annual benefits of about US$ 9 million—and an annual consumer benefit was about US$ 4 million. The results demonstrate that IAR4D adoption yields a rate of return of 38% and a benefit : cost ratio of 44 to 1. The average annual present producer surplus and present consumer surplus for millet were US$ 4.1 million and US$ 1.6 million, respectively, in the Sudan Savanna. The results further demonstrate that in millet production, the IAR4D approach yields a rate of return of 29% and a benefit : cost ratio of 20 to 1. In the same vein, the average annual present producer surplus and present consumer surplus, with respect to sorghum, are US$ 6.7 million and US$ 2.7 million, respectively. Also, the IAR4D approach in sorghum production yields a rate of return of 35% and a benefit : cost ratio of 33 to 1.

In the Northern Guinea Savanna, the estimates obtained for annual present producer surplus and present consumer surplus due to IAR4D in maize production are US$ 3.1 million and US$ 1.7 million, respectively, while the rate of returns and benefit : cost ratio are 27% and 15 to 1, respectively. Similarly, in the same zone, the annual present producer surplus and present consumer surplus, with respect to rice, are about US$ 13 million and US$ 5 million, respectively, and the rate of return and benefit : cost ratio are 42% and 67 to 1, respectively. Also, the annual present producer surplus and present consumer surplus as regards sorghum are about US$ 3.2 million and US$ 1.3 million, respectively, with the rate of return and a benefit : cost ratio of 27% and 16 to 1, respectively.

The estimates obtained for the Sahel Savanna were also similar to what were obtained in the other agro-ecological zones. With respect to millet, the estimated annual present producer surplus and present consumer surplus are about US$ 7 million and US$ 3 million, respectively, and a rate of return of 35% and a benefit : cost ratio of 34 to 1. On the other hand, the annual present producer surplus and present consumer surplus with respect to sorghum are about US$ 2.6 million and US$ 1 million, respectively, with the rate of return of 24% and a benefit : cost ratio of 12 to 1. Similarly, the annual present producer surplus and present consumer surplus as regards groundnut are about US$ 6.1 million and US$ 2.4 million, respectively, and a rate of return of 33% and a benefit : cost ratio of 29 to 1.

In the Lake Kivu PLS, results of the potential economic surplus model showed that, Rwanda gains an estimated US$ 285 million—equivalent to US$ 8 million per year—from adoption of the IAR4D approach in pepper production.

Sorghum in the Democratic Republic of Congo generated through IAR4D estimated gains of US$ 391 million—equivalent to US$ 11.2 million per year. The average annual present producer surplus and present consumer surplus are US$ 8.2 million and US$ 3.3 million, respectively, with a rate of return of 37% and a benefit : cost ratio of 42 to 1. In Uganda, IAR4D generated an estimated gain of US$ 359 million—equivalent to about US$ 10.3 million per year. The average annual present producer surplus and present consumer surplus were about US$ 7.5 million and US$ 3.0 million, respectively. The benefits due to potato in Uganda are the highest in the country and higher than what obtains for the same crop in the other two countries in the PLS.

For the ZMM PLS, the results obtained from the analyses of the data suggest that if the technology had been available at the baseline year and priced appropriately so that it would be adopted comprehensively, farmer's benefits in the PLS would have been US$ 382 million in that year; this is shared as US$ 61.75 million to Zimbabwe, US$ 145 million to Mozambique and US$ 174 million to Malawi.

The project had a bigger impact on the poorest beneficiaries and could have much greater impact in the future because of the time-lagged effect of the productive asset acquisition. Thus, a follow-up study is needed to capture the longer-term effects of productive assets and other changes that farmers experienced as a result of participation in the IAR4D. This study was conducted at an early stage of the project and does not capture its time-lagged impacts, especially the long-term benefits of productive asset acquisition and rural infrastructure development.

Key issues that need to be addressed in scaling up this success story include, among others, better targeting of poor and vulnerable groups, especially women; finding sustainable methods of promoting development of rural financial services; and conscious inclusion of capacity building of IAR4D beneficiaries in efficient management of productive assets.

As regards appropriate targeting, recall that over the first two years that the project operated, the Gini coefficient of income for beneficiaries decreased by about 6%, compared with a decrease of 5% for the conventional sites and an increase for other categories of non-beneficiaries. This suggests that the project contributed to a reduction in income inequality, probably through targeting poor and vulnerable groups. Consistent with this, the project also succeeded in raising the value of productive assets of the poorest tercile more significantly than for the other terciles. The non-significance of the impact on income for the other two terciles suggests appropriate targeting of the poor and vulnerable groups.

# Annex 1. Methodological Framework for Impact Evaluation of the Sub-Saharan Africa Challenge Programme (SSA CP)

## Introduction

A broad-based consensus and coalition is building around the urgent need to meet the Millennium Development Goals (MGDs) of halving hunger and poverty by 2015, and the importance of the agricultural sector in meeting these goals. However, despite decades of investment in public policy and development intervention projects and programmes, especially agricultural development programmes, and evidence of high return from such investment, hunger and poverty continue to plague large areas of the developing world, especially sub-Saharan Africa (SSA) (Alene et al. 2007). According to Baker (2000), billions of dollars are being spent on development assistance each year, yet little is known about the actual impact of projects on the poor. This suggests that there is very low impact of these interventions on the majority of the inhabitants of the African continent. In effect, the value of the frequent and large investments by donor agencies is not reflected in any way.

## Need for Impact Evaluation

In recent times, donors have insisted on getting value for money invested in any development intervention. As such, the possibility of accessing funds (especially for agricultural or pro-poor projects) from donor agencies is subject to the inclusion of concise and detailed impact evaluation procedures in the planning cycle of the project in question. Baker (2000) submitted that though there are broad benefits from economic growth, investments in human capital, and the provision of safety nets for the poor, questions arise regarding a specific programme or project in a country, such as the following:

1. Is the intervention producing the intended benefits and what was the overall impact on the population?
2. Could the programme or project be better designed to achieve the intended outcomes?
3. Are the resources being spent efficiently?

These types of questions can only be answered through an impact evaluation, an approach that measures the outcomes of a programme intervention in isolation of other possible factors. However, these questions cannot be answered simply by listing the outcomes of a project. There may be other factors or events that are correlated with the outcomes, but that are not caused by the project. To ensure methodological rigour, an impact evaluation must estimate the counterfactual, that is, what would have happened had the project never taken place or what otherwise would have been true. For example, if a recent graduate of a funded agricultural extension training programme becomes employed, is it a direct result of the training programme or would that individual have found work anyway? To determine the counterfactual, it is necessary to net out the effect of the interventions from other factors, a somewhat complex task. This is accomplished through the use of comparison or control groups (those who do not participate in a programme or receive treatment/benefits), which are subsequently compared with the treatment group (individuals who do receive the intervention). Control groups are selected randomly from the same population as the programme participants, whereas the comparison group is more simply the group that does not receive the programme under investigation. Both the comparison and the control groups should resemble the treatment group in every way, the only difference between groups being programme participation.

Determining the counterfactual is at the core of evaluation design. This can be accomplished using several methods, which fall into two broad categories: experimental designs (randomized), and the quasi-experimental designs (non-randomized). It is, however, quite tricky to net out the programme impact from the counterfactual conditions that can be affected by history, selection bias, and contamination. Qualitative and participatory methods can also be used in addition to assess impact. These techniques often provide critical insights into beneficiaries' perspectives, the value of the programmes to the beneficiaries, the process that may have affected the outcomes, and a deeper interpretation of the results observed in quantitative analysis.

## Approaches to Impact Evaluation

Several approaches can be used to evaluate programmes (Baker 2000; Khandker et al. 2010). However, a comprehensive evaluation is defined in the literature as an evaluation that includes monitoring, process evaluation, cost-benefit evaluation, and impact evaluation. Yet each of these components is distinctively different. Monitoring efforts track the key indicators of progress over the course of a programme, as a basis on which to evaluate outcomes of the intervention. Monitoring a programme also enables continuous feedback on the status of programme implementation, identifying specific problems as they arise. There is also operational evaluation, which examines how effectively programmes were implemented and whether there are gaps between planned and realized outcomes. Impact evaluation studies whether the changes in well-being are indeed due to programme intervention, and not attributable to other factors. Process evaluation is concerned with how the programme operates and focuses on problems in service delivery. Cost-benefit or cost-effectiveness evaluations assess programme costs (monetary or non-monetary), in particular their relation to alternative uses of the same resources, and to the benefits being produced by the programme. In all, impact evaluation is intended to determine more broadly whether the programme had the desired effects on individuals, households, and institutions and whether those effects are attributable to the programme intervention. Impact evaluations can also explore unintended consequences, whether positive or negative, on beneficiaries. In summary, evaluation can be carried out in two major ways (Baker 2000): quantitative methods and qualitative methods. Quantitative evaluation can be either experimental or quasi-experimental in nature. In the qualitative method, impact evaluation is carried out with the intent to determine impact by reliance on something other than the counterfactual to make a causal inference (Mohr 1995). The focus instead is on understanding the process, behaviours, and conditions as they are perceived by the individuals or groups being studied (Valadez and Bamberger 1994). The strengths and weaknesses of each of these methods are discussed in Table A1.

**Table A1. Summary of the possible approaches to impact evaluation.**

| **Quantitative Approaches** | | | | |
|---|---|---|---|---|
| **Approach** | **Description** | **Benefits/ Advantages** | **Problems/ Disadvantages** | **Some examples of case studies** |
| **1. Experimental Designs** | | | | |
| **1.1 Experimental or Randomized Control Designs (Randomiza-tion)** | Selection into the treatment and control groups is random within some well-defined set of people. In this case, there should be no difference (in expectation) between the two groups besides the fact that the treatment group had access to the programme. (There can still be differences due to sampling error; the larger the size of the treatment and control samples, the lesser the likelihood of error.) | The main benefit/ advantage is in the simplicity of the technique in interpreting results: the programme impact on the outcome being evaluated can be measured by the difference between the means of the samples of the treatment group and the control group. | 1. Randomization may be unethical owing to the denial of benefits or services to otherwise eligible members of the population for the purpose of the study.<br>2. The scope of the programme may mean that there are no non-treatment groups, such as with a project or policy change that is broad in scope.<br>3. Individuals in control groups may change certain identifying characteristics during the experiment that could invalidate or contaminate the results. | 1. Kenyan textbooks evaluation, in which evaluators selected a random allocation of programme sites, administered a baseline survey, created control groups, and then administered the treatment, which in this case was the delivery of textbooks. |

*Continued...*

*Table A1 Continued...*

| Approach | Description | Benefits/ Advantages | Problems/ Disadvantages | Some examples of case studies |
|---|---|---|---|---|
| **2. Non Quasi-experimental designs** | | | | |
| **2.1 Matching methods or constructed controls** | Here, one tries to pick an ideal comparison that matches the treatment group from a larger survey. The most widely used type of matching is "propensity score matching", in which the comparison group is matched to the treatment group on the basis of a set of observed characteristics or by using the "propensity score" (predicted probability of participation given observed characteristics); the closer the propensity score, the better the match. A good comparison group comes from the same economic environment and was administered the same questionnaire by similarly trained interviewers as the treatment group. | These methods can draw on existing data sources and are thus often quicker and cheaper to implement, and they can be performed after a programme has been implemented, given sufficient existing data. | 1. The reliability of the results is often reduced as the methodology is less robust statistically.<br>2. The methods can be statistically complex.<br>3. There is a problem of selection bias. | 1. Farmer-field school programme in Peru (Godtland et al. 2004).<br>2. Trabajar workfare programme in Argentina (Jalan and Ravalion 2003) |
| **2.2 Double difference or difference-in-differences methods** | One compares a treatment and a comparison group (first difference) before and after a programme (second difference). Comparators should be dropped when propensity scores are used and if they have scores outside the range observed for the treatment group. | These methods can draw on existing data sources and are thus often quicker and cheaper to implement, and they can be performed after a programme has been implemented, given sufficient existing data. | 1. The reliability of the results is often reduced as the methodology is less robust statistically.<br>2. The methods can be statistically complex.<br>3. There is a problem of selection bias. | 1. Female school stipend programme in the Punjab province of Pakistan (Chaudhury and Parajuli 2006).<br>2. Impact of a development programme in a poor area on growth in household consumption (Jalan and Ravalion 1998a). |

| Approach | Description | Benefits/ Advantages | Problems/ Disadvantages | Some examples of case studies |
|---|---|---|---|---|
| **2.3 Instrumental variables or statistical control methods** | One uses one or more variables that matter to participation but not to outcomes given participation. This identifies the exogenous variation in outcomes attributable to the programme, recognizing that its placement is not random but purposive. The "instrumental variables" are first used to predict programme participation; then one sees how the outcome indicator varies with the predicted values. | These methods can draw on existing data sources and are thus often quicker and cheaper to implement, and they can be performed after a programme has been implemented, given sufficient existing data. | 1. The reliability of the results is often reduced as the methodology is less robust statistically.<br>2. The methods can be statistically complex.<br>3. There is a problem of selection bias. | 1. Food for Education programme I Bangladesh (Ravallion and Wodon 1998b).<br>2. Child health and nutrition on education outcomes in Ghana (Glewwe and Jacoby 1995).<br>3. Microfinance programme in Bangladesh (Pitt and Khandker 1998). |
| **2.4 Reflexive comparison methods** | A baseline survey of participants is done before the intervention and a follow-up survey is done after. The baseline provides the comparison group, and impact is measured by the change in outcome indicators before and after the intervention. | These methods can draw on existing data sources and are thus often quicker and cheaper to implement, and they can be performed after a programme has been implemented, given sufficient existing data. | 1. The reliability of the results is often reduced as the methodology is less robust statistically.<br>2. The methods can be statistically complex.<br>3. There is a problem of selection bias. | |
| **The Qualitative Approach** | The qualitative approach uses relatively open-ended methods during design, collection of data, and analysis. Qualitative data can also be quantified. Among the methods used in qualitative impact assessments are the techniques developed for rapid rural assessment, which rely on participants' knowledge of | 1. They are flexible.<br>2. Can be specifically tailored to the needs of the evaluation, using open-ended approaches.<br>3. Can be carried out quickly, using rapid techniques. | 1. Subjectivity is involved in the data collection.<br>2. Lack of statistical robustness.<br>3. Lack of comparison groups.<br>4. All these, given small samples. | 1. REDP microhydropower projects in Nepal (see description in Khandker et al. 2010).<br>2. The FONCODES project in Peru (Schady 1999; Paxson and Schady 2002). |

*Continued...*

*Table A1 Continued...*

| Approach | Description | Benefits/ Advantages | Problems/ Disadvantages | Some examples of case studies |
|---|---|---|---|---|
| | the conditions surrounding the project or programme being evaluated, or participatory evaluations in which stakeholders are involved in all stages of the evaluation—determining the objectives of the study, identifying and selecting indicators to be used, and participating in data collection and analysis. | 4. Can greatly enhance the findings of an impact evaluation, through providing a better understanding of stakeholders' perceptions and priorities and the conditions and processes that may have affected programme impact. | | |

## Approach of Impact Evaluation Employed in Analyzing the Impact of the IAR4D, and the Justification for its Adaptation

In choosing an impact evaluation methodology, given the variations in project types, evaluation questions, data availability, cost, time constraints, and country/region circumstances, each impact evaluation study will be different and will require some combination of appropriate methodologies, both quantitative and qualitative (Baker 2000). The evaluator must carefully explore the methodological options in designing the study, with the aim of producing the most robust results possible. Among the quantitative methods, experimental designs are considered the optimal approach and matched comparisons a second-best alternative. Other techniques, however, can also produce reliable results, particularly with a good evaluation design and high-quality data.

Given concerns with the implementation of the randomized evaluations, the approach (randomization) is still a perfect impact evaluation method in theory (Khandker et al. 2010). Thus, when a treatment cannot be randomized, the next best thing to do is to mimic randomization—that is, try to have an observational analogue of a randomized experiment. Thus, in this report we employed the matching method of evaluation—the *propensity score matching* (PSM) method—because counterfactuals and or control groups have been identified and developed during programme planning. These counterfactuals are as far as possible similar to the treatment groups in terms of observational characteristics. So from the separately large groups of available non-participants, individuals who are *observationally similar* to participants in terms of characteristics not affected by the programme are evident.

The PSM method constructs a comparison group by modeling the probability of participating in the programme on the basis of observed characteristics unaffected by the programme.

Participants are matched on the basis of this probability, or propensity score, to non-participants. The average treatment effect of the programme is then calculated as the mean difference in outcomes across these two groups (in the case of the SSA CP, the treatment is separately compared with each of the conventional and clean groups). On its own, PSM is useful when only observed characteristics are believed to affect the programme participation. This assumption hinges on the rules governing the targeting of the programme, as well as any factors driving self-selection of individuals or households into the programme. Ideally, if available, pre-programme baseline data on participants and non-participants can be used to calculate the propensity score, and to match the two groups on the basis of the propensity score.

The use of the PSM approach in this report is justified for the following reasons: (1) Baseline data on participants and non-participants, which were collected through the baseline survey of 2008, are available; (2) in effect, comparable groups are available for both the baseline and midline data; and therefore, (3) it is possible to achieve sufficient robustness in the results.

## Analytical framework for the evaluation of the IAR4D's impact

Evaluation approaches for development programmes have evolved considerably over the past two decades, spurred on by expanding research on impact evaluation and growing coordination across different research and policy institutions in designing programmes (Khandker et al. 2010). Comparing programme effects across different regions and countries is also receiving greater attention, as programmes target larger populations and become more ambitious in scope, and researchers acquire enough data to test specific policy questions across localities.

Intervention programmes are designed to reach certain goals and beneficiaries. Methods to understand whether such programmes actually work, as well as the level and nature of impacts on intended beneficiaries, are the objectives of this report. Effective impact evaluation should, therefore, be able to assess precisely the mechanism by which beneficiaries are responding to the intervention. These mechanisms can include links through markets or improved social networks, as well as tie-ins with other existing policies.

Effective development policy-making creates a need for reliable methods of assessing whether an intervention had (or is having) the intended effect (Essama-Nssah 2006). "There should, therefore, be an intimate relationship between effective policymaking and impact analysis". The goal of an intervention defines the metric by which to assess its effectiveness. Effective methods of evaluation produce reliable information on what works and why, and policymakers may use such information to modify or cancel ineffective programmes and thus make the most of limited resources (Grossman 1994).

Ex-ante evaluation predicts programme impacts using data before programme intervention, whereas ex-post evaluation examines outcomes after programmes have been implemented. Reflexive comparisons are a type of ex-post evaluation; they examine programme impacts through the difference in participant outcomes before and after programme implementation (or across participants and non-participants).

As mentioned earlier on, approaches for impact evaluation vary. The variations are based on whether qualitative or quantitative assessment is required at any particular point in time. The method that fulfils the primary objective of this report is quantitative and, as such, we summarise the Propensity Score-based methods employed in this report.

## Propensity Score Matching (PSM) method in theory

The assessment of the impact of a programme (or a development intervention) requires *a model of causal inference*. Holland (1986) specifies such a statistical model. He starts from the fundamental observation that the effect of a cause can be understood only in relation to another cause. This is the same idea underlying the economic principle of assessing the return to a resource employed in one activity relative to its opportunity cost (i.e., what it would have earned in the next best alternative use). Thus, we can assess the effect of a development intervention only if we know what would have happened without such an intervention. Consider a simple situation involving only two causes: programme participation versus non-participation. A statistical causal inference model applicable to such a case involves the following elements: (1) a population of units upon which causes or interventions may act (e.g., individuals, households, districts, firms or regions); (2) an assumption that each unit is potentially exposable to the causes; (3) an observable variable $\mathrm{d}$, indicating the cause to which a given unit is exposed (e.g., $\mathrm{d} = 1$ for exposure, and zero otherwise); (4) a set of variables representing pre-exposure attributes for each unit (some attributes may be observable, call them $\mathrm{x}$, and some not, call these ε); and (5) a variable, $\mathrm{y(d)}$, representing the potential response of a unit to exposure. In fact, $\mathrm{y}$ represents two variables standing for two potential responses: $\mathrm{y1}$ under exposure, and $\mathrm{y0}$ if there is no exposure.

In theory, the PSM approach tries to capture the effects of different observed covariates $\mathrm{X}$ on participation in a single propensity score or index. Then, outcomes of participating and non-particpating households with similar propensity scores are compared to obtain the programme effect. Households for which no match is found are dropped because no basis exists for comparison. PSM constructs a statistical comparison group that is based on a model of the probability of participating in the treatment T conditional on observed characteristics $\mathrm{X}$, or the propensity score: $\mathrm{P(X) = Pr\ (T{=}1/X)}$. Rosenbaum and Rubin (1983) show that, under certain assumptions, matching on $\mathrm{P(X)}$ is as good as matching on $\mathrm{X}$. The necessary assumptions for identification of the programme effect are: (a) conditional independence and (b) presence of common support. The treatment effect of the programme using these methods can either be represented as the average treatment effect (ATE) or the treatment effect on the treated (TOT). Typically, researchers and evaluators can ensure only internal as opposed to external validity of the sample, so only the TOT can be estimated. Weaker assumptions of conditional independence as well as common support apply to estimating the TOT.

**Assumption of conditional independence.** “Conditional independence” states that given a set of observable covariates $\mathrm{X}$ that are not affected by treatment, potential outcomes $\mathrm{Y}$ are independent of treatment assignment $\mathrm{T}$. If $Y_i^T$ represent outcomes for participants and $Y_i^C$ outcomes for non-participants, conditional independence implies

$$(Y_i^T, Y_i^C) \perp T_i \mid X_i \qquad (1)$$

This assumption is also called "unconfoundedness" (Rosenbaum and Rubin 1983), and it implies that uptake of the programme is based entirely on observed characteristics. To estimate the TOT as opposed to the ATE, a weaker assumption is needed:

$$Y_i^C \perp T_i \mid X_i \qquad (2)$$

Conditional independence is a strong assumption and is not a directly testable criterion; it depends on specific features of the programme itself. If unobserved characteristics determine programme participation, conditional independence will be violated, and PSM is not an appropriate method. Having a rich set of pre-programme data will help support the conditional independence assumption by allowing one to control for as many observed characteristics as might be affecting programme participation (assuming unobserved selection is limited). Alternatives when selection on observed characteristics exists, and thus conditional independence is violated, include the instrumental variable and double-difference methods.

**Assumption of common support.** A second assumption is the "common support" or "overlap condition": $0 < P(T_i = 1 \mid X_i) < 1$. This condition ensures that treatment observations have comparison observations "nearby" in the propensity score distribution (Heckman et al. 1999). Specifically, the effectiveness of PSM also depends on having a large and roughly equal number of participant and non-participant observations, so that a substantial region of common support can be found. For estimating the TOT, this assumption can be relaxed to $P(T_i = 1 \mid X_i) < 1$.

Treatment units will, therefore, have to be similar to non-treatment units in terms of observed characteristics unaffected by participation; thus, some non-treatment units may have to be dropped to ensure comparability. However, sometimes a non-random subset of the treatment sample may have to be dropped if similar comparison units do not exist (Ravallion 2008). This situation is more problematic because it creates a possible sampling bias in the treatment effect. Examining the characteristics of dropped units may be useful in interpreting potential bias in the estimated treatment effects. Heckman et al. (1997) encourage dropping treatment observations with weak common support. Only in the area of common support can inferences be made about causality.

**The TOT using PSM.** If conditional independence holds, and if there is sizeable overlap in $P(X)$ across participants and non-participants, the PSM estimator for the TOT can be specified as the mean difference in $Y$ over the common support, weighting the comparison units by the propensity score distribution of participants. A typical cross-section estimator can be specified as follows:

$$TOT_{PSM} = E_{P(X)|T=1}\{E[Y^T \mid T = 1, P(X)] - E[Y^C \mid T = 0, P(X)]\} \qquad (3)$$

More explicitly, with cross-sectional data and within the common support, the treatment effect can be written as follows (see Heckman et al. 1997; Smith and Todd 2001):

$$\Rightarrow TOT_{PSM} = \frac{1}{N_T}[\sum_{i \in T} Y_i^T - \sum_{j \in C} \omega(i,j) Y_j^C] \quad (4)$$

Where $N_T$ is the number of participants $i$ and $\omega(i,j)$ is the weight used to aggregate outcomes for the matched non-participants $j$.

## Data collection

Task forces assembled data to establish baselines, monitored IAR4D processes, monitored the generation and use of innovations and evaluated their impacts. These data were collected at several levels, namely plot, household, village, innovation platform and district. Task force data were pooled at the PLS level to obtain a PLS perspective and subsequently at the programme level to obtain a sub-Saharan African perspective (see Figure A1).

## Baseline surveys for Innovation Platform and community levels characteristics

Baseline surveys, field observations and focus group discussions were conducted to benchmark pre-treatment characteristics of innovation platforms, site characteristics and baseline levels of outcomes predicted under the IAR4D approach: number, variety and time to develop innovations; knowledge and behavioural outcomes (adoption, input supply, input demand, volume of sales); market outcomes (output supply and consumption demand); and productivity outcomes (yields, technical and allocative efficiency and profit) and impacts (incomes, livelihood assets and equity). Several indicators were used to measure outcomes. These differed with context. Questionnaires were designed for comparison within IP over time and across IPs.

To generate counterfactuals, surveys and field observations were conducted in the comparison sites and villages assigned to conventional and non-IAR4D-non-conventional treatments. Key players in the innovation systems, such as public and private agricultural researchers, extension, farmer leaders, traders, dealers, lenders and key informants, were interviewed to benchmark characteristics of innovation systems and baseline levels of outcomes that were similar to those for the IP sites.

## Baseline survey for household and village community characteristics

Baseline surveys, observations and focus group discussions were conducted to collect data on household-level and village community-level characteristics and behavioural, efficiency, environmental and welfare outcomes. Surveys tracked feedback, information diffusion, awareness and knowledge changes, adoption, and market effects of innovations and spillovers using the Miguel and Kremer (2004) approach and other methods.

## Evaluation surveys

Follow-up evaluation surveys and qualitative assessment studies were conducted in the third year (2010) to assess the implementation process; document all the intermediary steps of the

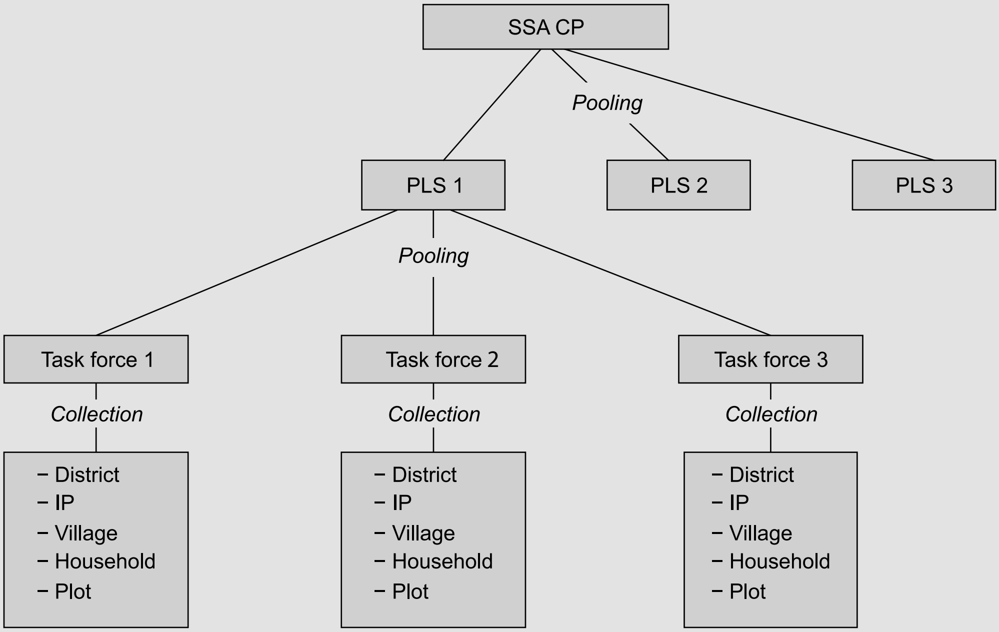



***Figure A1. The levels at which SSA CP data will be collected and pooled.***

research-to-impact pathway and conditioning factors; assess participants' subjective reactions to IAR4D; identify subgroups experiencing a greater or lesser impact than the sample as a whole; and measure changes in outcomes at the levels of the IP, household, community, and market. Follow-up surveys used the same indicators as were used in the baseline surveys to measure outcomes.

## Ex-Ante Estimation Method

To assess the potential economic benefits from adoption of the IAR4D approaches, we

(1) estimated the yield gains and the unit production cost reduction,

(2) examined the adoption pathway, and

(3) used an economic surplus model and information from the first three steps, along with secondary data, to evaluate the potential economic impacts of the IAR4D approach.

### Yield gains and unit cost reduction

Fora of stakeholders (consisting of farmers, agro-allied companies, seed producers) were established in 2004 in the the various PLS by FARA, in collaboration with the task forces in each PLS. Demonstration trials were conducted in collaboration with all the relevant stakeholders at the country levels. In each country, demonstrations were done in 4 districts. Five villages each were chosen from each of the districts. For easy comparison, equal numbers of villages were chosen from areas where there was some ARD and no ARD. All the farmers in the IAR4D and ARD have similar socio-economic characteristics, and they were exposed to the same improved crop and livestock varieties.

Using data obtained from on-farm trials, estimates of the yield gain and unit cost reduction effects of the IAR4D approach on crop production, for different countries, can be derived as follows: The priority crops were established. The estimation of the yield gains and cost reduction are as follows: (1) the average yield for adopters of the IAR4D option was estimated for the baseline, together with an estimate of the increase of the yield for the crops over the ARD approach.

However, the share of arable lands in each country was established from the available database. According to the database, total arable land for the country was obtained. Based on the importance of the crops in each country, the land allocated for the major crops were determined. The shares of land for respective crops were multiplied with the average yield of each crop in the respective countries, in order to get the total output for the crop.

### The adoption pathway

We used household survey data to project the adoption patterns of the IAR4D approach over time. We assumed that adoption started in 2006, and that nearly 10% of the sample households in the pilot villages adopted by the end of the year. Adoption picked up in 2008 and, by the end of 2009, about 30% of the households would have adopted. We assumed that by the end of 2010, about 50% would have adopted the IAR4D option. The estimates on adoption rates were used to extrapolate the ceiling adoption rates that can be expected across each of the countries. Since the household survey was undertaken in an area where adoption had occurred and was occurring, the percentage of farmers who would adopt the IAR4D approach in 2010 was assumed to be the ceiling rate of adoption—50%. However, it was assumed that for total coverage of each of the agro-ecological zones, the ceiling rate of adoption of the approach would only be reached in 15 years, as opposed to the 5 years it took the project villages in the pilot villages to reach this ceiling. In other words, it would take the whole ecological zone about 3 years to achieve the rate of adoption that the project villages achieved in one year. The adoption rates in the target villages from 2006 to 2010 and the assumed adoption lag for coverage of the each of the agro-ecological zones were used to estimate the parameters of the logistic function needed for predicting adoption rates into other areas in the zones from 2012 to 2035, as follows:

$$A_{it} = \frac{C_i}{1+e^{-(a+bt)}} \tag{1}$$

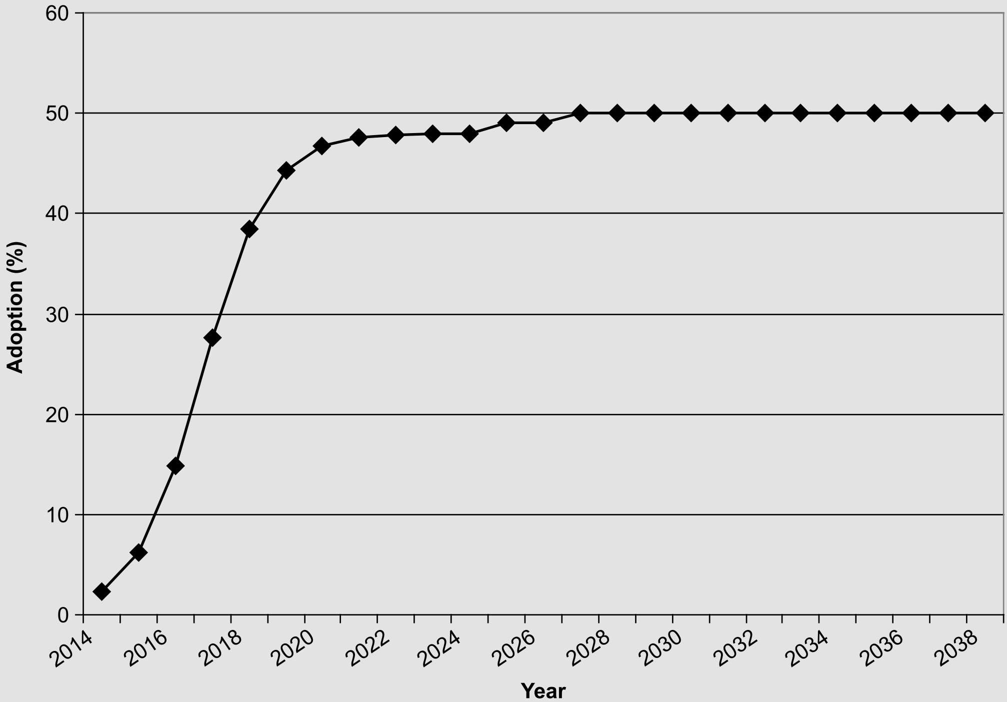
Where $A_{it}$ is the percentage adoption of the $i^{th}$ IAR4D approach in the $t^{th}$ year; $C_i$ is the adoption ceiling of the $i^{th}$ technology; $b$ is the rate of adoption; and $a$ is the constant intercept term. The adoption pathway for IAR4D option was predicted, using the following logistic function that was estimated using the survey data:

$$A_t = \frac{50}{(1 + e^{(4.1 - 1.1t)})} \qquad (2)$$

Supply shift. The unit cost reduction as a proportion of product price (discussed earlier) represents the maximum supply shift ($K$)—given 100% adoption—and translates into the actual annual supply shift ($Kt$) when multiplied by innovation adoption at time $t$ ($At$). That is, the annual supply shift is the product of cost reduction per tonne of output as a proportion of product price ($K$) and technology adoption at time $t$ ($At$). Indeed, the standard supply-and-demand diagram, demonstrating shifts in the supply curve due to adoption of a new technology, represents research benefits for one year. A successful research investment will yield benefits over a number of years. As the level of adoption increases, there will be further shifts in the supply curve, and corresponding changes in benefits (see Figure A2).

Estimating research benefits. The potential benefits of a technical intervention can be measured ex ante as well as ex post. A number of studies have applied the economic surplus model to estimate research benefits. The essence of the economic surplus model is that an improved technology, such as the IAR4D approach, reduces the cost of production of each kilogram of

***Figure A2: Projected adoption of the IAR4D option.***



output, leading to a shift in the supply curve to the right, an increase in the quantities supplied and traded, and a drop in prices in a competitive market. When this happens, although the selling price is reduced, small-holder producers may benefit from the reduced production costs and from selling larger quantities of the commodity produced at these lower costs, while consumers benefit from lower purchase prices. Two scenarios are always presented: the closed and the open economy models.

Assuming a closed economy model implies that the adoption of the IAR4D option increases the supply of crops. This study used a partial equilibrium, comparative static model of a closed and an open economy and a simple case of linear supply and demand with parallel shifts. A review of research benefits by participants revealed that most studies have used the assumption of linear supply and demand curve. However, when a parallel shift is used, the functional form is largely irrelevant, and the linear model provides a good approximation to the true (unknown) functional form of supply and demand.

A hypothetical case is illustrated in Figure A3. The supply of any given crop before the technical intervention of IARD approach is denoted by $S_0$. The demand for maize is denoted by $D$. The supply of maize shifts to $S_1$ following adoption, changing the equilibrium price and quantity before intervention from $P_0$ and $Q_0$ to a new equilibrium price and quantity, $P_1$ and $Q_1$. The change in consumer surplus is the area represented by $P_0ABP_1$, and the change in producer surplus is the area covered by $P_1BCD$. The change in total surplus is the sum of consumer and producer surpluses, which can be shown to be equal to $I_1I_0AB$.

In a closed economy, economic surplus measures can be derived using the following formulas: (1) economic surplus $(\mathrm{ES}) = P0Q0Kt(1{+}0.5Zt\eta)$; (2) consumer surplus $(\mathrm{CS}) = P0Q0Zt(1{+}0.5Zt\eta)$; and (3) producer surplus $(\mathrm{PS}) = (Kt-Zt)P0Q0(1{+}0.5Z\eta)$, where $Kt$ is the supply shift representing the product of cost reduction per tonne of output as a proportion of product price ($K$) and technology adoption at time $t$ ($At$), both of which have been presented and discussed earlier; $P0$ represents pre-research price (US$/tonne); $Q0$ is the quantity of commodity produced in tonnes; $\eta$ is the price elasticity of demand; and $Zt$ is the relative reduction in price at time $t$, which is calculated as $Zt = Kt\varepsilon/(\varepsilon+\eta)$, where $\varepsilon$ is the price elasticity of supply. Similarly, in a small open economy, change in economic surplus is equal to change in producer surplus and can be calculated as $\mathrm{ES} = \mathrm{PS} = PwQ0Kt(1{+}0.5Kt\varepsilon)$, where $P\mathrm{w}$ is the real world price.

**Prices and price elasticities.** In view of the fact that cereals are a highly tradable commodity in regional as well as international markets, however, the base model uses the open economy framework, and average international maize prices for 2006–2010, adjusted for shipping and insurance, were used in valuing the research benefits. The average real international crop prices were estimated based on FAOSTAT database. There is no reliable estimate of cereals and legumes supply elasticity for West Africa; in such situations, the price elasticity of cereals and legumes supply was assumed to be 1. Given that the crops are important staples for most households in West Africa, the (absolute) price elasticity of demand for cereals was assumed to be 0.4.

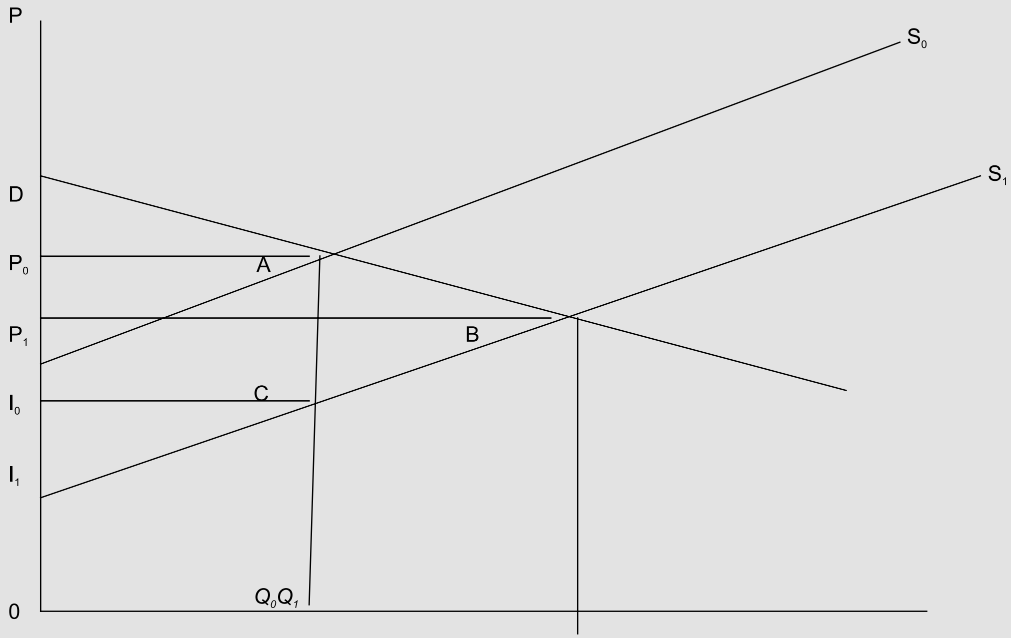



*Figure A3. Estimating changes in producer and consumer surplus.*

**Research and extension costs.** The total costs for research, development and extension of IAR4D approaches from 2004 to 2013 were obtained from the IAR4D approach management at FARA's IAR4D project management's office at the CIAT and the Institutes of Agriculture. The research costs included the annual salary of the IAR4D task force and other implementation teams; the annual operational expenses required to set up various IPs and sustaining them, as well as other costs involved to undertake the approach, including packaging and diffusion of the IAR4D option; and the annual overhead costs at the FARA. The annual extension cost associated with the large-scale dissemination of the approach in each country was estimated at US$ 1.5 million for each crop for the expected 15 years—from 2013 until the adoption of IAR4D approach reaches the ceiling of 50%.

## Data analysis: Estimation issues

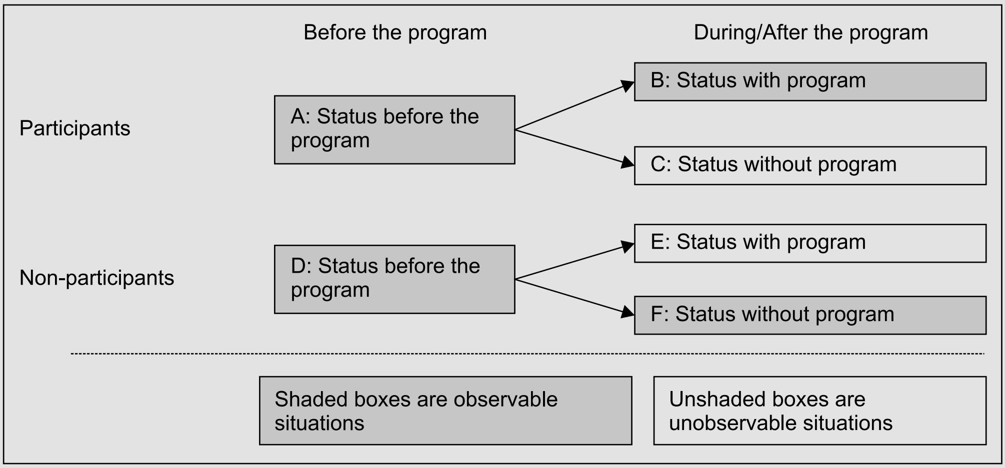
**Quantitative analysis.** The fundamental evaluation problem in estimating the effects of the IAR4D approach is the attribution problem and constructing counterfactuals. The counterfactuals, i.e., what would have happened to participants and non-participants without the programme, are never observed. In Figure A4, A and D and B and E can be observed but not C and F. How can B–C be estimated if there are no observations? An assumption often made is that E = F, that is, there is no self-selection among programme participants, scale effects, and spillovers. But the programme might affect prices in general, and there may be social and economic interaction effects on participants and non-participants. Potential spillovers and scale effects determine selection of with and without programme analytical approaches.

Estimation methods include

- Longitudinal comparisons of participants' outcomes (B–A), that is *before and after* treatment;
- Cross-sectional comparisons of participants' outcomes versus non-participants (B–E );
- Social experiments (B–E , with A = D and C = F); i.e., random assignment assures treatment households participating in the programme and non-treatment households are statistically equivalent. The key assumption is E = F, that is, there are no effects of the programme on non-participants;
- Difference-in-differences estimator ((B–E)–(A–D)), which accounts for fixed differences between outcomes of participants and non-participants;
- Matching (B–E|A(X)=D(X)); i.e., which compares outcomes of participants and non-participants who are similar in observed characteristics;
- Econometric methods (B-E|X), which account for impacts of observable and unobservable confounding factors (X) on outcomes; and
- Combinations of the above (e.g., difference-in-difference with econometrics: (B–E|X1)–(A–D|X0))

Longitudinal comparisons of participants' outcomes are most commonly used, in general. Differences between outcomes of project participants after and before the programme are used to measure the effects of the programme. Pre-programme data are used to impute the missing counterfactual outcomes for participants. An advantage of this before-after estimator, relative to other estimators, is that it can be implemented even when data are available only on participants (Todd 2006). However, longitudinal comparisons are likely to produce biased

***Figure A4. The problem of cause-and-effect attribution in assessing impact.***



estimates of treatment effects because they do not separate such effects from possible confounding factors (A≠C), such as bad or good weather, prices of crops, and long-term trends.

Cross-sectional comparisons of participating and non-participating households can improve attribution to treatments. They involve using data on a comparison group of non-participants to impute the counterfactual outcomes for participants. This estimator has the added advantage of demanding minimal data. However, its estimates will be biased if participants and non-participants are very different (presence of selection bias, e.g., poor and rich households) because the estimated cross-sectional differences will not be due only to the programme (A≠D, E≠C).

A social experiment overcomes selection bias by randomizing placement, so that all observational units within some defined set have the same probability ex-ante of receiving the treatment. If the treatment assignment is random and there is full compliance, i.e., focal villages in IAR4D sites cannot reject treatment and focal villages in the control sites and the non-IAR4D-non-conventional sites cannot implement IAR4D, then the assessment of the effect of IAR4D on outcomes such as yield at the household level can be done by taking the difference in means. Random assignment ensures that observed and unobserved characteristics of treatment and non-treatment households have the same distribution, i.e., A = D and C = F. Statistically B–C = E–F = F–C . Thus it resolves the problem of purposive placement based on unobserved factors. The assumption of no spillover and scale effects is very important for social experiments: large spillovers invalidate the purpose of counterfactuals.

The randomized design used for this study sought to minimize the spillover problem by assigning the IAR4D treatment at the district/local government areas/communes level, instead of the individual treatment (using control village communities at a sufficient distance from treatment village communities). However, randomized social experiments can alter the way the programme works in practice, because institutional and political factors may delay randomized assignment and randomization only yields mean outcomes for the counterfactuals, rather than distribution of outcomes and gainers and losers.

The difference-in-differences estimator nets out pre-project differences between participants and non-participants, such as initial differences in wealth, from the final difference. This measures the pure effect of the programme ((B–E)–(A–D)). The approach can be generalized to multiple periods. The double difference method has the advantage that it removes selection bias if the effects of selection bias are additive, and if the time invariant and outcomes are not affected by expectations of participation. Therefore, the method can be used even if purposive sampling is used to select households. But double differencing has several problems. These include selection bias that may not be time invariant, such as differential growth rates due to different initial endowments; and sensitivity to data quality, since measurement errors are more serious in comparing changes in variables than comparing levels.

Propensity score matching methods involve identifying a sample of comparator non-participants that are as similar as possible to participants in their predicted likelihood of participation and then comparing mean outcomes. For example, baseline data can be used to select for participants and non-participants who are similar in observed characteristics and to compare differences across space or over time. The strengths of the approach are reduced dependence on parametric assumptions and reduced bias from comparing non-comparable observations. The weaknesses are that only selection on observables is addressed and selection bias resulting from unobservables may still remain; reliance on the parametric model to predict participation; heavy reliance on extent and quality of data to predict participation; and difficulties finding comparable non-participants which, in turn, results in sample truncation. Other weaknesses with the approach are that it requires a larger baseline survey, since non-comparable participating and non-participating households are dropped and it affects the population for whom impacts are assessed.

Econometric methods account for predictive effects of other factors, that is, the impact of observable and unobservable confounding factors $(X)$ on outcomes across individuals or over time. Consequently, they are often used in studies of impacts. Econometric modeling has the advantage that it can control for selection biases by accounting for observable differences between programme participants and non-participants ("selection on observables" and "selection on unobservables"). But econometric approaches suffer from several problems:

- parametric approaches depend on valid parametric assumptions;
- non-parametric approaches rely on large sample size and good data;
- identification of suitable instrumental variables (IVs) is often difficult;
- estimation based on instrumental variables only evaluates impacts due to variation in IVs; and
- biases may result from comparing non-comparable observations.

Several developments have taken place in the econometric modeling literature on treatment effects that make it a powerful approach for analyzing evaluation problems addressed in this study (Heckman 2005), such as the following:

1. Development of an explicit framework for outcomes, measurements and choice of outcomes where the role of unobservables in creating selection problems and justifying estimators is modeled.
2. Extensions to analyze subjective evaluations of outcomes and using choice data to infer outcomes.
3. Extension to model *ex-ante* and *ex-post* realizations and evaluations of treatments, regret and anticipation by agents.
4. Development of models for identifying entire distributions of treatment effects (*ex ante* and *ex post*), rather than the mean parameters traditionally estimated by statisticians. These distributions can be used to determine the proportion of people who benefit from the treatment.
5. Identification of distributional criteria allowing for analysis of alternative social welfare criteria for outcome distributions comparing different treatment states.

6. Modeling of simultaneous causality that relaxes recursive frameworks and allows the analyses of social interactions, general equilibrium effects and scale-up effects.

We have thus far in this Annex highlighted the different quantitative methods employed in the study being reported. A cocktail of methods was employed as most suitable for our purpose, because no single evaluation method can claim to be ideal in all circumstances. Different combinations of quantitative approaches have also been shown in the past to improve robustness (Ravallion 2008):

- Multiple methods were used to increase confidence in conclusions, since each method has different strengths and weaknesses.
- Econometrics or propensity score matching was used with double-difference estimator to limit or account for effects of pre-project differences between participants and non-participants.
- Econometrics was employed with instrumental variable methods to address potential biases caused by selective participation.
- Propensity score matching was combined with econometrics to limit the sample analysed econometrically to comparable units.

**Qualitative assessment approaches.** Qualitative approaches were used, including impact pathway analysis, outcome mapping, participatory evaluation, and developmental evaluation. These approaches involve engaging partners and stakeholders to lay out their theory of change and hypotheses about how they expect impact to be achieved. Monitoring and evaluation were linked to the realization of the expected impacts. Monitoring and evaluation thus focussed on key essential factors to enable outcomes to take place, e.g., for a market network monitoring would include how household indicators change, as also would network indicators if these are part of the impact pathways.

The strengths of qualitative assessments include better understanding of processes by which impacts come about and stakeholders' perceptions. For example, farmers and traders can provide insights about which mechanisms are most important in generating impact. Furthermore, understanding programme details and processes is a precursor for understanding selection issues and identifying instrumental variables in econometric modeling. Because IAR4D is flexible and adaptive, qualitative approaches are important since they allow for adaptation over time. Qualitative approaches are especially useful for organizational learning and change, and for understanding the determinants and constraints of IAR4D adoption and diffusion.

Qualitative methods were used to reinforce quantitative methods, rather than to substitute them. While the quantitative methods addressed the question of what is the impact, their qualitative counterparts addressed the question of why and how impact is or is not being achieved. Qualitative methods enable better understanding of programme theory and context, which provides knowledge on what is working well and and what is not, thereby making assessment more relevant to decision makers. They can be used to establish how to apply lessons learnt elsewhere and, therefore, resolve the question of external validity.

Understanding the diffusion of information will be essential in assessing spillover effects, which may be revealed by quantitative baselines. However, combining this information with qualitative analysis will produce more solid conclusions. The main weakness with qualitative approaches is that, by themselves, they are unable to attribute impacts to interventions. Another problem is that they result in sampling and interviewer biases. For the reasons enumerated above, a combination of quantitative and qualitative evaluation approaches will continue to be proposed for monitoring, evaluation and impact assessment of IAR4D.

# References


Abadie, A. and G. Imbens, 2006. *On the failure of the bootstrap for matching estimators*. NBER Working Paper No. T 0325. Cambridge, MA: National Bureau of Economic Research. Also available at http://papers.ssrn.com/sol3/papers/cfm?abstract_id=912426

Alene, A. D., V. M. Manyong, E. F. Tollene and S. Abele, 2007. Setting strategic agricultural research priorities based on potential impacts on poverty reduction: the case of Nigeria. Pages 8-26 in *Assessing the impact of Agricultural Research on Rural livelihoods: Achievements, gaps and options* (Alene, A.D., V. M. Manyong, S. Abele and A. D. Sanogo, Eds). International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria.

Ayanwale, A.B., A. A. Akinola and V. A. Adeyemo, 2011. *Ex Ante Analysis of IAR4D Concept in the Sub-Saharan Africa Challenge Programmme (SSA CP) of the Forum for Agricultural Research in Africa (FARA)*. Internal Report submitted to FARA, November 2011. 177pp.

Baker, J. L., 2000. *Evaluating the impact of Development Projects on poverty: A handbook for practioners*. The International Bank for Reconstruction and Development/The World Bank, Washington, D.C., USA.

Becker, S. and A. Ichino, 2002. Estimation of Average Treatment Effects based on Propensity Scores. *The Stata Journal*, 2(4): 358-377.

Chaudhury, N. and D. Parajuli, 2006. *Conditional cash transfer and female schooling: The impact of the Female School Stipend Program on the public school enrollments in Punjab, Pakistan*. Policy Rsearch Working Paper 4102, The World Bank, Washington, D.C., USA.

Dehejia, R. and S. Wahba, 2002. Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics* 84(1): 151-161.

Duflo, E., S. Mullainathan, and M. Bertrand, 2004. How much should we trust difference-in-difference estimates. *Quarterly Journal of Economics* 119(1): 249-275.

Duflo, E., R. Glennerster, and M. Kremer, 2006. *Using randomization in Development Economics research: A toolkit.* Working Paper No. 138. Center for International Development, Harvard University, Cambridge, MA, USA.

Essama-Nssah, B., 2006. *Propensity Score Matching and Policy Impact Analysis: A Demonstration in Eviews*. Poverty Reduction Group (PRMPR). The World Bank, Washington, D.C., USA.

FAO, 2009. *The state of food insecurity in the World*. Food and Agricultural Organization of the United Nations, Rome, Italy. 125 pp.

FARA, 2009. *Sub-Saharan Africa Challenge Programme: Research Plan and Programme for Impact Assessment*. Forum for Agricultural Research in Africa, Accra, Ghana. 36 pp.

Glewwe, P. and H. Jacoby, 1995. An economic analysis of delayed Primary School enrollment in a low income country: The role of early childhood nutrition. *Review of Economic Statistics,* 77(1): 156-169.

Godtland, E., E. Sadoulet, A. de Janvry, R. Murgai and O. Ortiz, 2004. The impact of Farmer-Field Schools on knowledge and productivity: A study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change*, 52 (1): 129-158.

Grossman, J. B., 1994. Evaluating social policies: Principles and U.S. Experience. *The World Bank Research Observer* 9 (July): 159-180.

Heckman, J.J., 2005. *The Science Model of Causality*. Available at http://www.nuff.ox.ac.uk/users/nielsen/res/Heckman/Heckman_sci-causality.pdf

Heckman, J. J., H. Ichimura and P. Todd, 1997. Matching as an econometrics evaluation estimator: Evidence from evaluating a job training program. *Review of Economic Studies* 64(4): 605-654.

Heckman, J., H. Ichimura, J. Smith and P. Todd, 1998a. Characterising selection bias using experimental data. *Econometrica* 66: 1017-1099.

Heckman, J., H. Ichimura and P. Todd, 1998b. Matching as an econometric evaluation estimator. *Review of Economic Studies* 65(2): 261-294.

Heckman, J. J., R. LaLonde and J. Smith, 1999. The economics and econometrics of active labour programs. Pages 1865-2097 in *Handbook of Labour Economics*, Vol. 3 (O. Ashenfeller and D. Card, Eds). North-Holland, Amsterdam, the Netherlands.

Holland, P. W., 1986. Statistics and causal inferences. *Journal of the American Statistical Association.* 81 (396): 945-950.

Jalan, J. and M. Ravallion, 1998a. *Income gains from Workfare: Estimates for Argentina's TRABAJAR Program using Matching Methods*. Development Research Group, The World Bank, Washington, D.C., USA.

Jalan, J. and M. Ravallion, 1998b. *Transfer benefits from Workfare: A Matching Estimate for Argentina*. Development Research Group, The World Bank, Washington, D.C., USA.

Jalan, J. and M. Ravallion, 2003. Estimating the benefit incidence of an antipoverty program by Propensity-Score Matching. *Journal of Business and Economic Statistics*, 21 (1): 19-30.

Khandker, S. R., G. B. Koolwal and H. A. Samad, 2010. *Handbook on Impact Evaluation: Quantitative methods and practices*. The International Bank for Reconstruction and Development/The World Bank, Washington, D.C., USA.

Miguel, E., and M. Kremer, 2004. Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72 (1): 159-218.

Mohr, L. B., 1995. *Impact Analysis for Program Evaluation*, 2nd ed. Sage Publications, Thousand Oaks, California, USA.

Paxson, C. and N. Schady, 2002. The allocation and impact of social funds: Spending on School infrastructure in Peru. *World Bank Economic Review*, 16 (2): 297-319.

Pitt, M. and S. Khandker, 1998. The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? *Journal of Political Economy,* 106 (5): 958-996.

Ravallion, M., 2005. *Evaluating anti-poverty programs*. Policy Research Working Paper No. 3625: The World Bank, Washington, D.C., USA.

Ravallion, M., 2008. Evaluating anti-poverty programs. Pages 3787-3846 in *Handbook of Development Economics*, Vol. 4 (Schultz, T.P. and J. Strauss, Eds). North-Holland, Amsterdam, the Netherlands.

Ravallion, M. and Q. Wodon, 1998. *Evaluating a targeted social program when placement is decentralized.* Policy Research Working Paper 1945. The World Bank, Washington, D.C., USA.

Rosenbaum, P. R. and D. B. Rubin, 1983. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39: 33-38.

Rosenbaum, P. R. and D. B. Rubin, 1985. The central role of the propensity score in observational studies for causal effects. *Biometrica* 701: 41–55.

Schady, N., 1999. *Seeking Votes: The political economy of expenditures by the Peruvian Social Fund (FONCODES), 1991-1995*. Policy Research Working Paper 2166, The World Bank, Washington, D.C., USA.

Smith, J. and P. Todd, 2001. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125(1-2): 305-353.

Todd, P.E., 2006. Evaluating Social Programs with endogenous program placement and selection of the treated. Pages 3847-3894 in *Handbook of Agricultural Economics*, Vol. 4 (T. P. Schultz and J. A. Strauss, Eds). North-Holland, Amsterdam, the Netherlands.

Valadez, J. and M. Bamberger (Eds), 1994. *Monitoring and evaluating social programs in developing countries.* Economic Development Institute, The World Bank, Washington, D.C., USA.

White, H. and K. Chalak, 2006. *Parametric and nonparametric estimation of covariate-conditioned average effects.* Discussion Paper, Dept. of Economics, University of California at San Diago, La Jolla, CA, USA. 26 pp.

World Bank, 2008. *World Development Report*. Agriculture for Development, The World Bank, Washington, D.C., USA.

# Acronyms and abbreviations

| | |
|---|---|
| ABU | Ahmedu Bello University, Nigeria |
| ARD | Agricultural Research and Development |
| ARMTI | Agricultural and Rural Management Training Institute (Ilorin, Nigeria) |
| ASARECA | Association for Strengthening Agricultural Research in East and Central Africa |
| ATE | Average treatment effect |
| ATT | Average treatment effect on the treated |
| AVRDC | Asian Vegetable Research and Development Center |
| CCARDESA | Centre for Coordination of Agricultural Research and Development for Southern Africa |
| CGIAR | Consultative Group on International Agricultural Research |
| CIA | Conditional Independence Assumption |
| CIAT | International Center for Tropical Agriculture |
| CIMMYT | International Maize and Wheat Improvement Center |
| CORAF/ WECARD | Conseil Ouest et Centre Africain pour la Recherche et le Développement Agricoles/West and Central African Council for Agricultural Research and Development |
| CSC | Common Support Condition |
| DARS | Department of Agricultural Research Services |
| DFID UK | Department for International Development United Kingdom |
| DDM | Double-Difference Method |
| EC | European Commission |
| FARA | Forum for Agricultural Research in Africa |
| GIS | Geographical Information System |
| IAR | Institute of Agricultural Research |
| IARC | International agricultural research centre |
| IAR4D | Integrated Agricultural Research for Development |
| ICRAF | International Center for Agroforestry Research |
| IFAD | International Fund for Agricultural Development |
| IFDC | International Fertilizer Development Center |

| | |
|---|---|
| IFPRI | International Food Policy Research Institute |
| IIAM | Instituto de Investigação Agraria de Moçambique |
| IITA | International Institute of Tropical Agriculture |
| INRAN | Institut National de Recherches Agronomiques du Niger |
| ISAR | Institut des Sciences Agronomiques du Rwanda |
| IP | innovation platform |
| IPG | international public goods |
| KKM | Kano-Katsina-Maradi (PLS) |
| KM | Kernel-based Matching method |
| KTARDA | Katsina State Agricultural and Rural Development Programme/Ministry of Agriculture |
| LK | Lake Kivu (PLS) |
| MDG | Millenium Development Goals |
| MTP | medium-term plan |
| NARS | national agricultural research system(s) |
| NGO | non-governmental organization |
| NNM | Nearest Neighbor-based Matching method |
| No. | number |
| PLAR | participatory learning and action research |
| PLS | Pilot Learning Site |
| PM&E | participatory monitoring and evaluation |
| PSM | Propensity Score Matching method |
| RAB | Rwanda Agricultural Board |
| RM | Radius Matching method |
| RPG | regional public goods |
| SC | Science Council of the CGIAR |
| OFECSA | Soil Fertility Consortium for Southern Africa |
| SRO | sub-regional research organization |
| SSA | Sub-Saharan Africa |
| SSA CP | Sub-Saharan Africa Challenge Programme |
| TOT | Treatment effect on the treated |
| TSBF-CIAT | Tropical Soil Biology and Fertility Institute of CIAT |
| USAID | US Agency for International Development |
| WOFAN | Women Farmers Association of Nigeria |
| ZMM | Zimbabwe-Mozambique-Malawi (PLS) |

## SSA CP Key Support Institution

FARA

CGIAR

European Commission

## SSA CP Donors

NORWEGIAN MINISTRY OF FOREIGN AFFAIRS

REPVBBLICA ITALIANA

WORLD BANK

IFAD
INTERNATIONAL FUND FOR AGRICULTURAL DEVELOPMENT

Government of the Netherlands

DFID Department for International Development

Ministry of Foreign Affairs of Denmark
DANIDA International Development Cooperation

## SSA CP Task Force Institutions

IFDC

IITA
Research to Nourish Africa

INRAN

RWANDA RAB AGRICULTURE BOARD

MAKERERE UNIVERSITY

Bioversity International

CIAT
Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture

Soil Fertility Consortium for Southern Africa
SOFECSA
CIMMYT

## SSA CP Coordinating Institutions

CCARDESA

ASARECA
Transforming Agriculture
for Improved Livelihoods

CORAF/WECARD

## SSA CP Partner Institutions

ILRI
INTERNATIONAL
LIVESTOCK RESEARCH
INSTITUTE

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

World Agroforestry Centre
TRANSFORMING LIVES AND LANDSCAPES

National
University of
Rwanda
nur
www.nur.ac.rw

AHMADU BELLO UNIVERSITY

AVRDC
The World Vegetable Center

INERA

University Of Zimbabwe
Knowledge - Diligence - Integrity

ICRISAT
Science with a human face

## About FARA

FARA is the Forum for Agricultural Research in Africa, the apex organization bringing together and forming coalitions of major stakeholders in agricultural research and development in Africa.

FARA is the technical arm of the African Union Commission (AUC) on rural economy and agricultural development and the lead agency of the AU's New Partnership for Africa's Development (NEPAD) to implement the fourth pillar of the Comprehensive African Agricultural Development Programme (CAADP), involving agricultural research, technology dissemination and uptake.

**FARA's vision**: reduced poverty in Africa as a result of sustainable broad-based agricultural growth and improved livelihoods, particularly of smallholder and pastoral enterprises.

**FARA's mission**: creation of broad-based improvements in agricultural productivity, competitiveness and markets by supporting Africa's sub-regional organizations (SROs) in strengthening capacity for agricultural innovation.

**FARA's Value Proposition:** to provide a strategic platform to foster continental and global networking that reinforces the capacities of Africa's national agricultural research systems and sub-regional organizations.

FARA will make this contribution by achieving its *Specific Objective* of sustainable improvements to broad-based agricultural productivity, competitiveness and markets.

Key to this is the delivery of five *Results,* which respond to the priorities expressed by FARA's clients. These are:

1. Establishment of appropriate institutional and organizational arrangements for regional agricultural research and development.
2. Broad-based stakeholders provided access to the knowledge and technology necessary for innovation.
3. Development of strategic decision-making options for policy, institutions and markets.
4. Development of human and institutional capacity for innovation.
5. Support provided for platforms for agricultural innovation.

FARA will deliver these results by supporting the SROs through these Networking Support Functions (NSFs):
NSF1/3. Advocacy and policy
NSF2. Access to knowledge and technologies
NSF4. Capacity strengthening
NSF5. Partnerships and strategic alliances

FARA's donors are the African Development Bank (AfDB), the Canadian International Development Agency (CIDA), the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), the Danish International Development Agency (DANIDA), the Department for International Development (DFID), the European Commission (EC), the International Development Research Centre (IDRC), the Syngenta Foundation, the United States Department of Agriculture (USDA), the World Bank and the Governments of Italy and the Netherlands.

FARA

**Forum for Agricultural Research in Africa**
12 Anmeda Street, Roman Ridge, PMB CT 173, Accra, Ghana
Telephone: +233 302 772823 / 302 779421; Fax: +233 302 773676 / Email: info@fara-africa.org
www.fara-africa.org