

Design and Implementation of Impact Evaluation of Market Systems Development (MSD) Projects:

The Case of InovAgro

Hosaena Ghebru¹, William Grant² and
Jenny Smart¹

¹ International Food Policy Research Institute - IFPRI

² Development Alternatives Incorporated – DAI

1-1-2021

Abstract

This paper identifies the lessons learned concerning the design and implementation of impact evaluation of Market Systems Development (MSD) programs using the case of Innovation for Agribusiness (InovAgro) program, a program which has been in effect from 2010 to 2020, that endeavors to transform critical market systems, both from a value chain perspective as well as in cross cutting services in Northern Mozambique's agricultural sector. This report offers observations concerning challenges faced and lessons learned in terms of both the collaborative design and execution of the project and its impact evaluation. The report gives special consideration to the broader systemic changes taking place within the project, as well as offers scrutiny of the timeframe required to achieve the desired broad-based results, both at the household and system levels.

Acknowledgements

This work was undertaken in collaboration with Development Alternatives Incorporated (DAI) and is part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). Funding support for this study was provided by the Swiss Agency for Development and Cooperation (SDC). This paper has not gone through IFPRI's standard peer-review procedure. The opinions expressed here belong to the authors, and do not necessarily reflect those of SDC, PIM, IFPRI, DAI or the CGIAR.

TABLE OF CONTENTS

I.	Conceptual Framework on Market System Development	4
II.	Market System Development Interventions and the InovAgro Project Theory of Change	6
	Overview, and theory of change of the InovAgro project	6
	Review of other major Market System Development interventions in developing countries ...	8
III.	Academic literature on M4P intervention evaluations	11
	1. Quantitative evaluations	13
	2. Qualitative evaluations	14
	3. The mixed methods approach	15
IV.	Empirical challenges: Lessons learned from the evaluation point of view	17
	1. Ethical issues:	17
	2. Challenges associated with implementational/operational issues:	18
	3. Challenges associated with level of impact (scope):	19
V.	Operational challenges: Lessons learned from implementation point of view	21
	1. Macro-economic factors and political economy	21
	2. Weather and Climate	23
	3. Global issues – policies in other countries impacting Mozambican farmers	23
	4. Meso-level institutional challenges	23
	5. Micro-economic (firm-level) challenges	24
	Adaptation in implementation	24
	Lessons learned from implementation	25
VI.	Recommendations for future evaluations of MSD programs	27
	1. Recommendations from the evaluator’s point of view	27
	2. Recommendations from the Implementer’s point of view	28
VII.	References	30

I. Conceptual Framework on Market System Development

The M4P (Making Markets Work for the Poor) approach focuses on improving the functioning of market systems in order to make these systems function well, in a sustainable manner, and in a fashion that improves the welfare of the poor. The premise behind this approach is that the poor participate in market systems every day,

and making the market systems more efficient for them is a critical element in poverty reduction. In particular, in the agricultural sector — on which the vast majority of the poor rely in many developing countries—enhancing the functioning of markets for the outputs of primary production, for inputs into production, and for processed products is crucial for addressing poverty. Identifying whether interventions that follow the M4P approach work, in what respects they work and in what ways they can yet improve, is a fundamental yet understated dimension in M4P activities.

The names Market Systems Development (MSD) and Markets for the Poor (M4P) are synonyms that are used by the existing literature to communicate the same concept. The concept of M4P and details of the approach are well stated by Tschumi and Hagan (2008). ‘Making Markets Work for the Poor’ is an approach to development support that was advocated and supported by the Swiss Agency for Development and Co-operation (SDC) and the Department for International Development (DfID). As it is stated by Tschumi and Hagan (2008), “M4P is an approach to developing market systems so that they function more effectively, sustainably and beneficially for poor people, building their capacities and offering them the opportunity to enhance their lives.” The rationale behind the M4P approach emanates from a position that market system has a role in reducing poverty. The agencies argue that the lessons learned from a range of development experiences and the success of M4P in practice have led them to adopt the approach. M4P works to improve the market systems in such a way that they can work for the poor, such as helping public investment catalyze private investment. Improving such market systems in the agriculture sector is especially crucial as the largest segment of the poor population in developing countries rely on agriculture for their livelihood. M4P gives emphasis to the challenges in agricultural input and output markets. In the InovAgro project, the goal of pro-poor market systems development is sought by transforming public investment into private engagement, namely between input companies and farmers.

This paper identifies the lessons learned concerning the impact pathways of the InovAgro Market System Development Project, an M4P program that focuses on critical market systems, both from a value chain perspective as well as in cross cutting services in Northern Mozambique’s agricultural sector. The implementation of the program is led by the Development Alternatives, Inc. (DAI), while the evaluation of the intervention is led by the International Food Policy Research Institute (IFPRI). Note that this report does *not* aim to serve the purpose of an evaluation of the program’s impact.¹ Rather, as authored together by actors from IFPRI and DAI, it offers observations concerning what worked well and what did not work well in terms of the collaborative

¹ The results of the mid-line and end-line impact evaluations are the subject of separate reports, respectively, the first of which was released in 2019 (Ghebru, Smart and Mogues).

design and execution of the project and its impact evaluation. The InovAgro project endeavors to transform systems for greater long-term sustainable growth and impact on the poor. In this review of challenges faced and lessons learned, we give special consideration to the broader systemic changes taking place within the project and among other things, as well as offering scrutiny of the timeframe required to achieve the desired broad-based results, both at the household and system levels.

The report proceeds as follows: section II gives an overview of the theory of change of the InovAgro project and other market system development interventions. Section III reviews the academic literature on MSD intervention evaluations. Section IV highlights the empirical challenges faced by the study and the resulting lessons learned from an evaluation point of view, while section V outlines the operational challenges faced and resulting lessons learned from the implementation point of view. Section VI offers recommendations for monitoring and future evaluations of MSD programs from both evaluation and implementation points of view and concludes with closing observations.

II. Market System Development Interventions and the InovAgro Project Theory of Change

Overview, and theory of change of the InovAgro project

The InovAgro program consists of three phases – a design and three year pilot phase from the end of 2010 to 2013 (InovAgro I), a four year expansion phase building off of the pilot (InovAgro II) from 2014-2017, and a three year wrap-up phase (2018-2020). The overall objective of the SDC InovAgro project is to increase income and improve economic security for poor men and women small-scale farmers in Northern Mozambique through improved agricultural productivity and the development of high potential value chains. An MSD program has a standard theory of change: it focuses its attention on the system level issues, in order to get the system to perform more effectively to meet the needs of the target beneficiaries to have improved access and growth. The project does not deliver the services directly to the target beneficiaries, since the project recognizes that it will not be around once the funding starts, so the market system must continue to drive the change. So most of the impact on the target beneficiaries, the rural poor, will be delivered from their engagement with the actors in the market system.

The project has a carefully articulated theory of change, but which does not completely follow the normal logframe input-output-outcome sequence. Figure 2 greatly simplifies the sequence of change, but demonstrates how the interventions undertaken under InovAgro — the activities at the Inputs level in the graphical illustration of this theory of change — would ultimately result in greater impact such as increased farmer income. The pathway from the activities to impact includes several levels of outputs and outcomes, including intermediate outcomes that cannot be reflected in a standard logframe. For example, changes in behavior by output buyers - Outcome Level - will lead to strengthened linkages with farmers (Output Level) will lead to increased participation by farmers in value chains to increase their incomes—the Outcome Level (Figure 2).

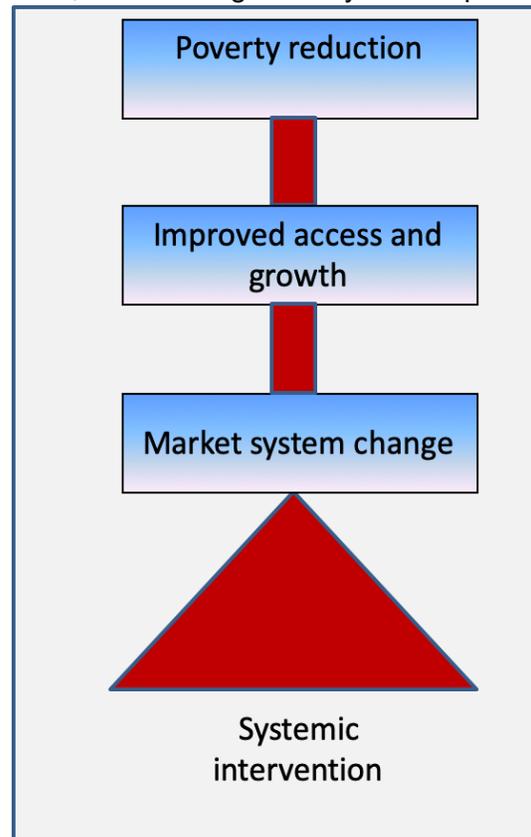
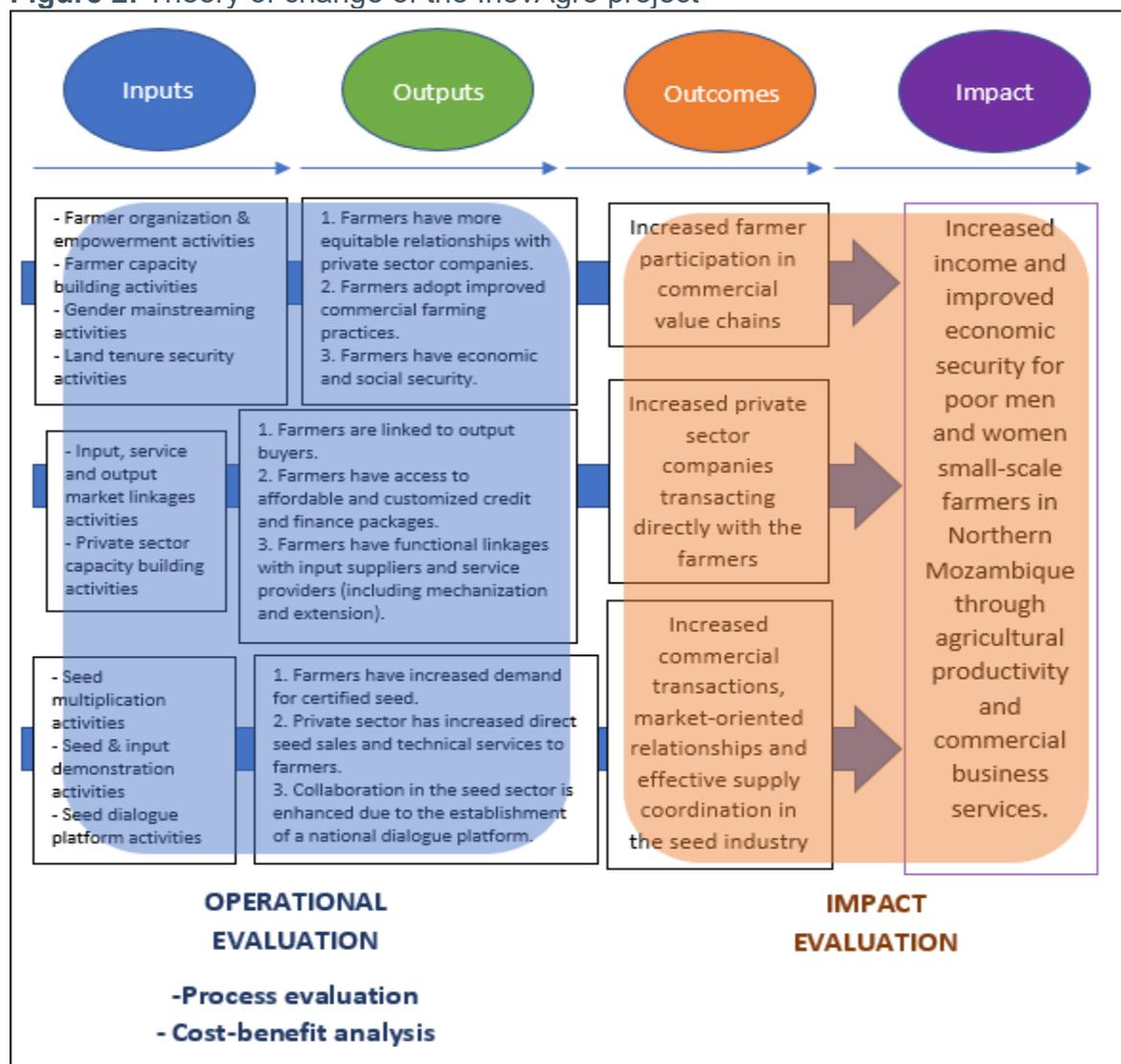


Figure 1: Generic MSD Theory of Change

Outcome Level - will lead to strengthened linkages with farmers (Output Level) will lead to increased participation by farmers in value chains to increase their incomes—the Outcome Level (Figure 2).

Figure 2: Theory of change of the InovAgro project



Source: Adapted from DAI (2013).

The impact evaluation traces the project's theory of change, by judicious identification of variables to measure and analyze. This contributes not only in identifying the impact of InovAgro, but also in shedding empirical light on the mechanisms through which such impact occurs, quantifying both impacts and outcomes. Impact evaluation of the outcomes to impact is a process that is complementary to the operational evaluation of tracing inputs, or project activities, into project outputs. Monitoring and evaluation, cost-benefit analysis and process evaluations using administrative data and qualitative assessments are all valuable tools used to conduct such operational evaluations. Operational evaluation and impact evaluation analyses are not mutually exclusive one to another and rather, combined, do the best job at offering the full story.

One method of operational evaluation that hints at the attempt to ascertain outcomes and systemic level impact, albeit not via a quantitative impact evaluation, is assessing the maturity level of the project activities in achieving a level of systemic change. The Initiative for Global

Development has leveraged the theory of the “tipping point” to gauge the level of adoption at which an innovation is considered to then be on a path towards eventual self-sustenance. According to the theory, this point is reached when roughly 16% of a population have adopted an innovation, given that such a level of adoption indicates that activity plans have become clearly defined and “sticky”, thus leading to the buy-in, consistency and strategy of stakeholders, funding, and/or monitoring which helps programs have what it takes to take off and build to scale (Gifford et al 2016).

Review of other major Market System Development interventions in developing countries

Although many interventions have been implemented using M4P approaches, there are fewer impact evaluations on M4P projects, and less still, that fulfill the academic-criteria of a standard impact evaluation. In this section of the report, we review other major Market System Development interventions in developing countries. Then, in section III that follows, we look at how the evaluation of these studies fail to meet the academic criteria of an impact evaluation. Finally, in sub-sections III.1 – III.3, we proceed to elaborate the academic criteria of a rigorous impact evaluation as undertaken by IFPRI in relation to InovAgro project.

Several interventions have been implemented using M4P with the objective of improving the livelihoods of the poor. In this section, we list M4D or M4P rural development interventions similar to InovAgro. For illustration purposes, we have presented three interventions that have been selected by Tschumi and Hagan (2008) on the basis of their performance, diversity of contexts, and being subjected to external reviews. In addition, we have included five more M4P interventions (4-8) in order to present a broader scope of M4P interventions².

1. *FinMark Trust - Financial services in South Africa*: The goal of FinMark in this intervention was to make financial service markets work more effectively for the poor. To this end, the project targeted low-reach financial service providers among low income groups and carried out the following: built a shared view of the industry’s future; supported service innovation; developed information services; improved regulatory process; and stimulated consumer education. By doing so, the project contributed to a 7.1 million increase in ‘banked’ population; new information services; stronger public roles; and a better

² We note that given poverty is gendered, it would have been interesting to know the gendered impact of these M4P interventions on poor women and men, however gendered results were not presented for these studies in as far as what was included in the accounts that we reviewed and summarized here, with the exception of the *MSME Strengthening Project in Cambodia*. For this project, gendered results of training and activity attendance are offered, and the simple take-away made that project staff should have recruited more women participants for achieving better gender balances in several areas. De Wildt’s report also shares that nursery trainings in the Katalyst – Fish Pond project in Bangladesh included information on crosscutting gender issues, and the second phase of the project set out to place specific emphasis on increasing outreach along gender lines and in an explicitly pro-poor way (2007). In a much more holistic approach, Jones and Oakeley explore a gendered approach in M4P by exploring the opportunities and challenges that are presented by a market-oriented perspective on women’s empowerment in their 2012 article. They suggest an approach for explicitly incorporating women’s economic empowerment in the M4P framework, going beyond the scope adopted by IFPRI’s evaluation design of InovAgro.

innovation environment (Tschumi and Hagan 2008).

2. *FIT-SEMA - Mass media (radio) in Uganda*: The goal of FIT-SEMA in this intervention was to improve the quality and relevance of commercial radio programming on a sustainable basis. To this end, the project targeted a small number of radio stations and worked to improve program innovation and quality as well as stronger market functions more widely (audience research, journalism etc.). By doing so, 25 stations were able to offer 50 new programs and thereby reached 7 million additional listeners (ibid).
3. *Katalyst - Vegetables project in Bangladesh*: The goal of this intervention was to improve the productivity of vegetable farming on a sustainable basis. To this end, the project provided support to input suppliers to develop a retailer training program and increase retailers' incentives and capacity to offer helpful advice to farmers. In this way, the project developed a new supply chain business model that functioned to serve and improve performance among 1 million farmers (ibid).
4. *Social Marketing of Insecticide - Treated Nets (SMITN) Project in Tanzania*: The goal of this intervention was to improve the use of Mosquito nets among the poor. The project promoted the distribution and sale of untreated mosquito nets (i.e. a dip-it-yourself insecticide treatment kit) through a network of retail distributors as well as through NGOs and public sector health facilities. Initially nets were sold at subsidized prices, but over time, the market price of domestically produced nets fell. By the end of the project, social marketing prices were slightly higher than the commercial sector prices, bed-net coverage had increased substantially in all the project areas and whereas socio-economic differentials within the population played a larger role in net ownership initially, by the end of the project, bednets were being employed much more evenly across income levels (Mills et al 2002).
5. *KINET project Vouchers for Mosquito Nets in Tanzania*: Another mosquito-net project in Tanzania, the goal of this intervention was also to improve the use of Mosquito nets among the poor. In this case, the project provided a 17 percent discount on the price of nets and promoted their sales at maternal and child health (MCH) clinics. The discount was provided in the form of voucher and could be used as partial-payment for a net purchased from any retailer. "The shopkeeper returned the redeemed vouchers to the wholesaler in exchange for the full value of the voucher plus a small margin. The wholesaler in turn returned vouchers to the project in exchange for the full value of the voucher plus an additional small margin." Although the redemption rate of the vouchers was very high (97 percent), awareness of the scheme was low (less than 50 percent). The study also indicated that vouchers were more likely to be used in the least poor, than in the most poor, households (ibid).
6. *Katalyst – Fish Pond project in Bangladesh*: The goal of Katalyst in this intervention was to increase the output of pond fish cultivation among poor fishermen. The project intervention targeted improved access to inputs, access to advice, and the role of sector associations. To address the problem of access to inputs, Katalyst provided technical

assistance to the Faridpur Fisheries Association (FFA) to lobby for the development of a large-scale physical market for fingerlings (this included Katalyst conducting a feasibility study as well as contributing management input for the market's construction). Similarly, Katalyst provided technical assistance to help build the organizational capacity and effectiveness of fishery associations, first by offering a vision workshop and then offering further one-on-one advice. Moreover, Katalyst supported the delivery of a three-day training programme for nursery-owners in how to offer better information and advice to fish farmers. This helped farmers, nurseries and hawkers increase productivity and revenues (de Wildt 2007).

7. *Micro-Small and Medium-sized Enterprise (MSME) Strengthening Project in Cambodia:* The goal of this project was to improve entrepreneurship and competitiveness for micro, small and medium enterprises (MSME) in selected value chains and target provinces by enabling improvements to businesses and the business environment. To this end, the project engaged MSMEs in the three primary value chains of swine, aquaculture, and masonry (brick and tile making), in six provinces of Cambodia. By partnering directly with the value chain stakeholders, the project assisted rural MSME entrepreneurs by promoting new ways of thinking and acting among local business people in order to improve opportunities and income. By doing so, the project has helped improve the productivity and business performance of these enterprises and generated wider beneficial impacts at the value chain level as well (USAID 2008).

The fresh vegetable component of the Growth-Oriented Microenterprise Development Program (GMED) in India: The goal of this intervention was to link small-scale farmers to competitive value chains. To this end, GMED followed a parallel approach that included technical assistance and the establishment of vertical relationships. The technical assistance offered involved the establishment of demonstration farms and carefully developed technical packages used in the training of extension personnel who worked with lead farmers who, themselves, became farmer-to-farmer trainers and buyers of other farmers' upgraded outputs. The vertical relationships were thus mutually beneficial and linked small-scale farmers to corporate buyers. A standard impact evaluation of the study indicated that the project was largely successful in achieving the outcome of increased farmer awareness of process and product upgrading (Dunn et al. 2011).

III. Academic literature on M4P intervention evaluations

Although many interventions have been implemented using M4P approaches, there are fewer impact evaluations on M4P projects, and less still, that fulfill the academic-criteria of a standard impact evaluation. Ruffer and Wach (2013) also conclude that M4P evaluations have been generally weak in terms of considering the systemic, sustainable changes in market systems; the use of theories of change; triangulation practices; data quality; consistency in units; and consideration of unintended negative effects. In this regard, however, we would like to mention that there are many reviews of activities of M4P projects that have been undertaken either by external consultants—hired by donors or implementing agencies—or reviews undertaken by the implementing agencies themselves. These studies vary by way of quality, though most fail to meet the academic criteria of an impact evaluation, which we proceed to elaborate in the following sub-sections.

Table 1: Summary of projects/evaluations reviewed

		Study design (parameters) - impact evaluation			
	Reviewed MSD programs/projects	If Impact Evaluation (IE) was conducted ?	If IE was independent or project-sponsored (by implementing firm)?	If IE was Qualitative or Quantitative or both	If IE was micro (household), meso (firm) or macro (market/systemic) level
1	<i>FinMark Trust - Financial services in South Africa</i>	NO			
2	<i>FIT-SEMA - Mass media (radio) in Uganda</i>	NO			
3	<i>Katalyst - Vegetables project in Bangladesh</i>	NO			
4	<i>Social Marketing of Insecticide - Treated Nets (SMITN) Project in Tanzania</i>	NO			
5	<i>KINET project Vouchers for Mosquito Nets in Tanzania</i>	YES	Only internal result tracking	Quantitative and Qualitative	Household level
6	<i>Katalyst - Fish Pond project in Bangladesh</i>	YES	Only internal result tracking	Quantitative	Household level only
7	<i>Micro-Small and Medium-sized Enterprise (MSME) Strengthening Project in Cambodia</i>	YES	Independent	Quantitative	Client level only
8	<i>The fresh vegetable component of the Growth-Oriented Microenterprise Development Program (GMED) in India:</i>	YES	Independent	Quantitative and Qualitative	Household level only

As shown on Table 1 above, our review has found that none of the 8 MSD projects (evaluations) applied a methodology to evaluate impact at market/systemic level. Interestingly, of the 8 projects/evaluations, only 2 (MSME and GMED) conducted a proper independent impact evaluation while another two (KINET AND Katalyst – Fish pond project) only involved internal (project-sponsored) result tracking.

The impact evaluation of the fresh vegetable component of the India GMED project is the only standard impact evaluation study of the M4P evaluation studies listed in the previous section. This evaluation was conducted by an independent party. The study examined project outputs, outcomes and impacts by means of four major research components: i) an assessment made in 2006 to define the causal model and design the evaluation; ii) a longitudinal, quasi-experimental survey, in which 712 farmers in treatment and control villages were interviewed in 2007 and 2009; iii) a process evaluation to assess the implementation history of the project; and iv) a qualitative field study to understand the pathways by which change occurred. The study specifically measured the impact indicators of household income, enterprise income, jobs and poverty. It also considered creation of sustainable vertical linkages explicitly (Dunn et al. 2011).

The Cambodia Micro-Small and Medium-sized Enterprise (MSME) project was also evaluated by an independent party. The evaluation study was conducted by randomly drawing a sample from a population of 889 MSME clients. The study used a before/after randomized sample comparison to identify the impact of the project. The study tried to have a control group by selecting a random sample of swine and fish producer enterprises that had not been former clients of the project's implementing agency. The study concludes that the project has had major impacts on improving the productivity and business performance of these enterprises, with wider beneficial impacts at the value chain level. It specifically measured the impact indicators of household and enterprise income (USAID 2008). Although this study fails to fulfil the academic standard of impact evaluation study, it can be considered as one of the M4P evaluation studies that partially meet the criteria of a standard impact evaluation study.

Similarly, the "impact" of the Katalyst-Pond Fish project in Bangladesh has been assessed by the UK Springfield Centre for Business in Development, commissioned by Katalyst Bangladesh. The study used a survey comprising 563 respondents. Although the study stated that productivity has improved among half of the farmers (13,600), half of the nurseries (400) and the majority of hawkers (4400) (de Wildt 2007), it is not clear how it reached this conclusion. The study also admitted that the impact of Katalyst's interventions need to be seen within the wider picture of continuous growth, rising consumer prices and targeted development projects.

The operation of the KINET project that distributed vouchers for mosquito nets in Tanzania was assessed through household surveys, a series of focus group discussions, and key informant interviews. The focus group discussions were with community leaders and parents of children under 5 years while the key informant's interviews were with maternal and child health (MCH) clinics staff and net retailers. The study concluded that voucher redemption rate was 97 percent though only 43 percent of mothers were aware of the voucher program and only 12 percent of

mothers said they had used a voucher to purchase a net. Moreover, the study found that vouchers were more likely to be used in the least poor than in the poorest households (Tami et al. 2005). However, the study doesn't have any counterfactual that shows what the purchases of nets would have been in the absence of the intervention.

Palmer and Mills (2003) compare an alternative private contractual mechanism for the delivery of primary care with two models where the public sector block purchased services on behalf of state patients. The study found that the introduction of the private sector contract has helped to reduce the two key hindrances—the specification and monitoring of contracts—that are common in the public sector. However, this study cannot be taken as an impact evaluation study as it lacks the core tools to ensure attributability that an impact evaluation requires. Overall, there are three broader approaches in conducting impact evaluations of MSD programs: namely, quantitative, qualitative, and mixed methods.

1. Quantitative evaluations

InovAgro's approach is based on the rationale that the private sector creates pro-poor growth opportunities. The overall objective of InovAgro is to sustainably improve the revenue of small female and male farmers through achieving the following outcomes: increased participation of farmers in commercial value chains of pigeon pea, soya bean, sesame and groundnut; increased commercial transactions between private sector companies and farmers in input, output, finance, and mechanized services markets; and, finally, an increased market oriented relationship, effective supply coordination and commercial transaction in the seed industry. In other words, the ultimate objective of the intervention is to reduce poverty by making markets better work for the poor, especially to the small farmer.

Thus, the key policy questions presiding to the intervention are: does InovAgro achieve these outcomes in such a way that the revenue and economic security of the small farmers increases during the life of the project and continues to increase even further for them and other farmers in the future? Is it worth replicating and expanding InovAgro to new geographic areas? Is it worth widening the InovAgro's approach and M4P interventions into different value chains? In order to answer such questions, standard impact evaluation of the intervention is important. And another related question of equal importance, is: what is the timeframe required to deliver changes at the farmer level, since the changes must be driven by the system?

Evaluations need to be scientific, by undergoing and passing formal peer-review, in other words by being published in peer-reviewed academic journal, in order to be credible and relevant in guiding policy directions. Similarly, the impacts of M4P interventions need to be measured scientifically so that the approach can ascertain itself in development practice. Above all, development approaches need to be tested scientifically for the effective and efficient allocation of the limited resources and for the improvement of the livelihoods of poor people.

From SDC's perspective, the objective of the impact evaluation study is to provide tangible and credible evidence using scientific approaches about the contribution of the InovAgro M4P approach (as opposed to direct delivery approach) to improving the lives of the beneficiaries and

leaving behind systems that will ensure the continued improvement, and ability to allow even more people than in the target beneficiaries to continue to improve. However, the objective of the impact evaluation study is not only to provide tangible and credible evidences about the contribution of the M4P approach to the lives of the beneficiaries, but to see the possibility of replicating the project in other regions and countries in an even more effective manner. Above all, quantitative impact evaluation provides credible evidence to the Government of Mozambique—especially to the Ministry of Agriculture—that can lead to the formulation of new agricultural policies and strategies that are aligned with a market development approach. Moreover, the experience obtained in the evaluation serves to get valuable lessons to strengthen future M4P projects' monitoring and evaluation systems.

Randomized control trials (RCTs) are considered the 'gold standard' for addressing the type of question that is at the heart of this study: What difference did the intervention make to the development outcomes considered, and how large is the difference that it made? The quantitative impact evaluation research design employed in the study is further detailed in the forthcoming midline Impact Evaluation report. RCTs are also typically used to assess direct impact on a narrowly defined beneficiary population. Therefore, relying on an RCT or quasi-experimental design alone to evaluate a market systems intervention will not offer the full extent of impact, since the goals of M4P are to influence systemic-level change in the market.

2. Qualitative evaluations

There is also a critical role that rigorously undertaken qualitative analysis provides in a study such as this one, where the intended goals of the intervention are at the system level. Specifically, there are four main areas in which qualitative approaches are a strong complement to quantitative methods and allow for a much more in-depth descriptive analysis:

- a. *Going beyond the “what and how much”: Addressing the “how”*: While quantitative methods are designed to do a good job of asking the first order question—what and how much, such as “what was the impact of the intervention”—they are less suited to ask the second order but still important question—the ‘how’, such as “how and through which mechanisms did the project achieve these outcomes”, as this is critical to understanding whether the benefits will last and continue to expand after the project? This is where qualitative techniques can make an important contribution to a study, especially in the context of the quantitative approaches including RCT (White 2013). In the case of InovAgro, qualitative investigation helps tease out the *pathways and mechanisms through which* InovAgro’s multifaceted support is strengthening the market participation and economic standing of beneficiary farmers. Analysis of the qualitative data gives insights on the impact pathways and mechanisms of InovAgro and allows an indicative empirical validation of or refinement on the theory of change underpinning the study. The “how” question investigated also includes an exploration of the political economy of agricultural markets, as part of an effort to understand the role that public policies, and political drivers of these policies, affect within the functioning of the local markets that are the focus of InovAgro.
- b. *Establishing external validity*: Quantitative tools, such as those described above, are powerful

for establishing internal validity of the study's findings. That is, they can ensure that they identify changes that were, in fact, truly contributed to by the project, as opposed to changes that may correlate in time, space, or some other way with the intervention activities but are not causally linked to the project. Internal validity is not easy to achieve, and weakly designed quantitative or qualitative work frequently fail this test. External validity, on the other hand, refers to the extent to which findings about causal impacts of an intervention or project in a particular area would also be applicable in other areas or contexts. It is, then, apparent that without solid internal validity, no external validity can ever be hoped for. However, of course having the former—which requires strong experimental or quasi-experimental methods—by no means automatically implies the latter. This is an area where qualitative endeavors contribute: Careful analysis (thorough understanding) of the process and pathways through which, and the context in which, InovAgro support is able to improve the functioning of agricultural markets, will shed light on the circumstances in which this kind of support is able to be successful in other geographical, agricultural, etc. contexts, and the conditions that are likely to render impact difficult in these other contexts.

- c. Exploring impact on overarching outcomes: Fundamentally, econometric and other strong quantitative techniques rely, among other things, on being able to examine variation across many units (e.g. farmers, firms, land plots, etc.) in terms of characteristics of these units (e.g. productivity, commercialization, income, profit, etc.). Some outcomes of interest however cannot be captured, not even proxied, at the multi-unit level. In M4P for example, there are a few outcomes of interest concerning the functioning of the overall market system—such as indicators of local market concentration—that are not well represented with indicators that can be measured at the level of the farmer, processor, retail input supplier, and so forth.
- d. Informing the quantitative data collection: The insights gleaned from qualitative fieldwork and analysis is, precisely because of its complementary nature, of extremely high use in informing the quantitative endline data collection, by allowing the team to further refine the choice of variables to be included in the endline questionnaire. The qualitative empirical data help uncover phenomena, previously unanticipated potential outcomes, and potential intermediate effects, which can then be included in the endline survey to be quantitatively examined.

Appropriate for the purposes of the qualitative-component analysis is a method referred to as process tracing (Beach and Pederson 2019), which is used to both test as well as build different elements of theories of change. While the quantitative analysis uncovers causal relationships—e.g. between the InovAgro intervention and outcomes of interest—process tracing is appropriate to uncover mechanisms of change. The key empirical tools that are used to employ this method are key informant interviews and focus group interviews.

3. *The mixed methods approach*

IFPRI employed a mixed methods approach of data collection and analysis in the evaluation of the InovAgro project. Mixed methods approaches are encouraged in impact assessment as they are complementary to one another (White 2013). Multiple quantitative methods also allow for comparing and testing of the sensitivity of results of one approach against the other. As referenced

earlier in this section, much more research has been done on the impact pathways of the M4P approach from a qualitative lens than that which has commented on the challenges and shortcomings of measuring impact pathways using standard quantitative impact evaluation approaches at the household level. For this reason, the rest of this report seeks to address this literature gap by expounding upon the lessons learned and recommendations identified by the InovAgro implementing and evaluating organizations in having sought to measure and achieve intervention outcomes that were targeted to improve conditions at the micro/farmer/household level.

IV. Empirical challenges: Lessons learned from the evaluation point of view

The on-going debate on the selection of appropriate impact evaluation methods for MSD projects continues within the development research and evaluation community, and indeed also within practitioner and policymaker circles (see, for example, Cohen and Easterly, 2009). This section discusses the multiple conceptual and methodological challenges in impact evaluation of MSD programs referring to randomized and non-randomized methods.

As a structure to the discussion, we focus on three key challenges in impact evaluation of such programs, namely:

- (1) Ethical issues
- (2) Challenges associated with implementation (operationalization) of programs
- (3) Challenges associated with level of impact (scope)

1. Ethical issues:

While using a randomized control trial is considered the gold-standard quantitative method for impact evaluation, one of the greatest challenges facing project evaluators in choosing to use it, both in the context of this evaluation and others, is the thorny ethical consideration of how one justifies the randomization process in practice: that is, the designating of some groups as treatment groups and others as the control groups. This can present difficulties particularly when it comes to evaluators needing to respond to the reactions of both beneficiary and non-beneficiary affected stakeholders. The experimental nature of a study's design must be handled delicately, as participants and others can have a hard time understanding the reasons for the perceived inequities in the roll-out of project interventions, can potentially grow resentful because of this, and many find that they simply cannot morally support the design framework given the disparities of how benefits are allocated across groups. Policy makers find the process of random allocation to be unacceptable in many cases as well (e.g. Bamberger and White, 2007; Ravallion, 2008). Interventions are often intended to be targeted to specific groups and this poses direct concerns to policy makers: Randomization might very well limit outreach at the same time as often being viewed as unethical when taken into consideration in light of the target population's pressing needs. Several mitigating measures were taken by the IFPRI impact evaluation team to address the associated ethical issues, namely: (i) after consultations with the implementing partner (DAI) and all stakeholders involved, the original study design of RCT at household level revised and adjusted to administer the randomization at the location (cluster or community). Thorough dialogue with the implementing firm and other involved stakeholders (including SDAE) made it possible to adopt such level of randomization as it involved less/no ethical issues. This is so since potential budget constraints and/or limited administrative/operational capacity of the local governments, donors and/or implementing agencies mean wider coverage of areas for the intervention (zone of influence) very difficult (impossible) which required gradual scale up and, hence, needing to randomize (especially when such programs are, indeed, implemented as part of a phased-in design).

2. Challenges associated with implementational/operational issues:

Another family of challenges which threaten the validity of a given impact evaluation method involve issues associated with the actual implementation of the MSD programs after the impact evaluation design is complete. Such issues include:

- (a) Selection bias issues
- (b) Contagion/contamination (treatment diffusion)/spillover effects
- (c) Historical phenomena/effects

a. Selection bias issues

Self-selection arises when farmers themselves decide whether or not to participate in MSD programs, often due to differences in resource endowments. Thus, the selection bias emanates from the fact that treated individuals may be systematically different from the non-treated for reasons other than the treatment status. In MSD projects, driving systemic change means driving behavioral change at the farmer level. This means that the project must have the most interested partners (farmers and service providers) participating in the project interventions who are the innovators and early adopters to drive the adoption and uptake of new innovations, to get the market system going and stimulate copying. The endogeneity problem may arise because the program may target farmers with specific characteristics (innovative farmers, community leaders, smallholder farmers, market-oriented farmers, poor farmers or relatively wealthy farmers). These may result in overestimating or underestimating the impacts of the MSD program. This is more pronounced in the pilot InovAgro impact evaluation study that the initial participants in MSD programs have a higher risk profile and are more willing to try new things during the baseline may on average have a better asset base compared to non-participants (perhaps, due to liquidity constraints (Ghebru et al, FORTHCOMING). To filter out such potential adverse effects due to selection bias, the InovAgro impact evaluation has adopted a propensity score matching technique to conduct a similar analysis on sub-sample of beneficiary and non-beneficiary households which are comparable characteristics (matched based on observable household and parcel level characteristics) at baseline (before program launch). Such matching method does not only help to filter out the potential biases in findings due to self-selection issues but also enables to investigate the robustness (sensitivity) of the impact evaluation results to self-selection biases.

b. Contagion/contamination issues

Once the design for impact evaluations of MSD programs is finalized and beneficiary and control groups are established, issues may arise in which the groups which are not supposed to be exposed to (or receiving) certain benefits (the control groups) are in fact benefiting from an intervention. This may arise either by directly receiving the benefits from the intervention (eg. in the case of InovAgro partners chasing the market and spreading the intervention to areas that were designated as controls), or by receiving similar benefits from other organizations who are not direct partners of InovAgro, but may be delivering similar services. The latter could also be the case due to local government budget shifts (e.g., similar interventions being financed via the SDAE(s)).the InovAgro impact evaluation has addressed such potential adverse effects on the

credibility of the impact evaluation by complementing the household panel survey data with a geo-referenced census data of all market actors (including SDAE interventions) operational in the study area as this helped to identify the zone of influence (household exposure) of these market actors.

c. Historical phenomena/effects

Historical phenomena effects are caused when behavioral change or adjustments by farmers and/or market actors are caused due to external events other than the intervention, occurring during the same time period as the intervention. In the case of Mozambique (InovAgro), the rapid devaluation of currency starting in January 2015 and the flooding in early 2015 are two noteworthy cases in this regard, with implications on the validity of impact evaluation designs for similar MSD programs, and more so in the former case, as devaluation might have direct effect on cost of imported agricultural inputs. Such effects could be due to a possible lower adoption of modern farm practices caused by inflated prices for imported inputs (devaluation effect) and/or slowing down of farmers' participation in high-value crops and shift towards lower-risk crops (natural disaster effect). As a result, the program effect could be understated if such external effects are not filtered out. The InovAgro impact evaluation has addressed such potential effects by adopting a double-difference econometric approach that utilized before-after data of not only beneficiary groups but also non-beneficiary groups. Hence, the validity of the InovAgro impact evaluation is maintained as such historical phenomena (effects) are expected to have similar effects on both intended beneficiary and non-beneficiary households.

3. Challenges associated with level of impact (scope):

a. Beneficiary versus institutional/systemic level effects

Broadly, intended outcomes of MSD projects can be categorized into two: beneficiary (primary) level and institutional (systemic) level effects. Beneficiary (primary) level effects are the direct impacts that an intervention seeks to achieve on targeted beneficiaries, such as increasing farm productivity and incomes for a target group or facilitating access to services by intended beneficiaries. However, most MSD programs (if not all) are primarily set out to produce outcomes that are to the benefit of the entire sector, not just for a few firms (more of an institutional or systemic level changes). It is, hence, always a challenge in designing and implementing impact evaluation methods that are sufficiently equipped to document impacts at both levels – i.e., micro as well as macro level impacts. In addition to the mitigative measures taken to address the study design issues related to micro (household) level effect, attempts have been made to document the potential macro (market/systemic) level effects. Hence, the InovAgro impact evaluation design utilized a more comprehensive approach by complementing the three-cohort household panel survey with a unique geo-referenced census data (GIS data) of all market actors operational in the two study districts which includes data on locations of the service provided, date of establishment, affiliations with upstream market actors at national and international level. The InovAgro impact evaluation report by Ghebru et al (forthcoming) details of the methodology and approaches adopted to account for the market/systemic level effects

b. Short-term versus long-term effects

Depending on what level of effect one tries to evaluate, for MSD interventions, primary-level effects (such as access to inputs, behavioral changes, etc.) may emerge quickly. In other cases, some effects may take much longer to become manifest, and change over time (such as income, poverty or overall welfare). MSD interventions are usually assumed to contribute to long-term development (income, poverty and overall systemic level effects). However, impact evaluations with a short evaluation time period, focusing on short-term or intermediate outcomes, often provide more useful and immediate information for policy- and decision-making but fail to document effects achieved in the longer term. Many of the beneficiary level impacts of interest from MSD interventions will only be evident in the longer-term, such as overall income and/or welfare impacts. Hence, searching for evidence of such effects too early (via scientific impact evaluation programs with short evaluation period) might mistakenly conclude that interventions have failed. Hence, it should be noted that ideal length of the impact evaluation period depends on the level of impact one seeks to document since the longer the project the more susceptible the impact evaluation design will be to challenges of spillover and contamination. This is evident from the InovAgro impact evaluation study where no statistically significant difference was recorded on some of short term/intermediary outcome indicators (access to market and agricultural information) almost 5 years after the project was implemented while the difference was evident during the mid-term assessment, i.e., 2 years following the project implementation (Ghebru et al, FORTHCOMING). This is recognized by the DCED which recommends collecting final impact data two years after the end of interventions.

c. Intended versus unintended effects

Another issue related to the selection and/or validity of impact evaluation methods of MSD projects is how well the Impact Evaluation (IE) method is equipped to deal with unintended effects. MSD programs usually are designed to have systemic level effects that are prone to unintended effects (positive or negative) on both intended beneficiaries and non-beneficiaries. For example, programs that aim to provide market-driven agricultural training, extension or credit services in a given area may lead to: (i) the government deciding to shift its budget allocations (cut its spending) for similar activities from beneficiary communities to control (non-beneficiary) communities (producing a negative unintended effect on beneficiary communities and an unintended positive effect on control communities); and, (ii) potential spillovers – i.e., households/farmers in control areas benefiting from market responses (input as well as output) due to their proximity to beneficiary communities. Both cases, unless addressed through appropriate impact evaluation methods, may lead to underestimation of the impact of MSD programs. The InovAgro impact evaluation has addressed the issue of documenting unintended effects by: (i) including longitudinal data not only from intended beneficiary households but also non-beneficiary households; ii) included in all the three-cohort household survey dataset (2014/15, 2016/17 and 2018/19) proxy variables that are believed to account for a host of possible unintended effects. This list includes: risk diversification (crop diversification); climate-friendly farming practice (fallowing); income diversification and migration (disaggregated data on employment and migration status by the head, spouse, female and youth member of the household); intra-household bargaining power and land rights (disaggregated data on access to credit, access to and control over land, possession of personal saving accounts, etc by head, spouse, any female and/or youth member of the household).

V. Operational challenges: Lessons learned from implementation point of view

All market systems projects operate in a complex environment which goes beyond the project's control and which adds challenges to delivering sound, sustainable, developmental results. Even with the best theory of change and well thought out causal model, projects must be flexible and able to adapt to changes in the broader environment which will affect the implementing environment. MSD projects must be able to take into account changes in macro-economic environment, the political economy, weather and climate (particularly true for an agricultural development project) crises and institutional responses, policy changes in other countries, national level regulatory policy (meso-level), and even at the corporate firm level. These changes can present new constraints, or sometimes opportunities, that will require strategic shifts in the implementation approach.

InovAgro has had to adapt to all of these externalities. We highlight a few of the most important external shocks that affected project implementation and requiring project adaptation.

1. Macro-economic factors and political economy

a. Government policy

There have been major shifts in the government policy on subsidized inputs. In 2011, the government was the main buyer of certified seeds in the country and then distributed them to the target farmers. This greatly affected the behavior of the three major seed companies producing certified seed in Mozambique, who focused their sales primarily on the government and very large farmers, with no attention to small farmers. However, in 2013, the government stopped purchases due to budget shortages and the three major seed companies all closed down because their major market had disappeared. The collapse of the government's ability to purchase and distribute seeds destroyed one model of seed companies in the country, setting the country supply of certified seed back several years, but opened up an opportunity for InovAgro to promote a much more dynamic, engaged, and competitive seed industry from a more diverse group of suppliers who now are targeting smallholder farmers directly. Without InovAgro's presence and its astute assessment of the opportunity created by the government's withdrawal, the small local seed companies would not have been able to develop as effectively as they have since 2015.

b. Currency devaluation

As all value chains are global, they compete with products from inside and outside the country. Therefore, trade conditions will be affected by the stability of the currency. Over a five-year time frame (2011-2016), the Metical ranged from 33 MZN, appreciated to 26.25 MZN and then depreciated to 71MZN to the US Dollar. This high fluctuation of the value of the currency has made it difficult for private sector seed companies (which initially have been importing certified seed) to plan, causing them to change their mix of strategies on importing vs local production. It has also affected the farmers' behavior and purchasing power, when faced with a doubling of the cost of imported seed, slowing adoption rates. This has allowed InovAgro to work with the seed

companies to help them develop alternative strategies for producing locally and segmenting their market.

From an impact assessment standpoint, the metical devaluation affected prices for the imported certified seed market (where the project was stimulating it), but not where it wasn't (the control districts) and imported certified seed was not for sale.

c. Political unrest

In addition to other issues, Northern Mozambique has seen recent conflicts flare up between RENAMO and FRELIMO, as well as attacks on private companies by insurgents in the far north of Cabo Delgado. Instability and conflict deter investment by private companies and prevent the transportation of goods and personnel across the conflict zones. Some partners who had worked with InovAgro in 2015 and 2016 curtailed their sales and marketing efforts into the North in 2017 following the conflicts. It has also hindered project staff ability to enter certain zones.

In response to this, InovAgro has developed stronger relationships with other NGOs and market actors from within the conflict zones who are able to circulate more freely and can represent private companies.

d. Political economy

Mozambique is in the process of major change with international investments seeking access to its natural resources, big companies seeking opportunities to establish monopoly rights in sectors, and opportunistic behavior. Two cases in particular elicited strong interventions from InovAgro: land grabs by large investing companies and an attempt to corner the pigeon pea market by a major trading company.

All rural land in Mozambique is owned by the government. While communities and farmers have traditional rights to use the land, the formal leasing process from the government can supersede traditional rights if they have not been formally registered. Following the displacement of a number of farmers by one major investor, InovAgro began a process for analyzing why farmers were not registering their land rights developing a market driven approach to land registration in Mozambique³.

The largest exporter of pigeon peas in Mozambique tried to capture the entire market through regulatory change. A push to require all local pigeon pea processing to be met before exports would be allowed, by the one company that had all of the pigeon pea processing capacity, would have reduced competition and lowered the price of pigeon peas to SHF.

³ [InovAgro](#): Applying a Market Systems Approach to stimulating land titling in Mozambique, June 2019.

2. Weather and Climate

During the implementation period under question, InovAgro's focal districts experienced a range of external weather-related challenges hampering implementation. In January 2015, El Nino related flooding wiped out many farmers' planting cycle, including the loss of all newly purchased certified seeds. At the same time, the flooding destroyed roads and bridges preventing project partners from reaching the zones that were affected.

The two most recent cyclones in 2019 (Idai and Kenneth) had significant impact on regions surrounding InovAgro's project districts. But most importantly, the broader systemic impacts from those two cyclones include destruction of seed production zones that were producing certified seed for sale in InovAgro districts. Challenges around availability of certified seeds for sale will likely be further exacerbated as NGOs and the Government are reacting to the shortage of seed in the areas affected by the storms and will likely purchase as much seed as is available now. These large bulk purchases of seed are easier for seed companies to meet than selling individually to smallholder farmers, which will distort the market for seeds and will likely decrease the availability of seed at planting season in November.

3. Global issues – policies in other countries impacting Mozambican farmers

Mozambique's exports of pigeon peas have been one of the major successes over the six year period from 2010 and 2016 as exports had soared from 30,000 MT to over 200,000 MT to India. However, in 2017, the Indian government instituted a new policy to stimulate pigeon pea production in India, leading to an increase in production of 2.2 million mt of pigeon pea (record production levels nearly double the preceding year), eliminating the need to import pigeon peas from anywhere. This wiped out most of Mozambique's exports and crashed the price of pigeon peas in Mozambique. As this had been a growing part of InovAgro's portfolio this impacted farmer profitability and behavior.

The project has responded by organizing information sharing workshops with the Ministry of Industry and Commerce and the International Growth Centre on the causes of the crash for pigeon peas but also on how to get the government to honor the trade agreement with India.

4. Meso-level institutional challenges

Government institutions and other donor projects can add to the challenges of stimulating systemic change.

Mozambican institutional capacity to certify seed has been a constraint. The seed companies faced the challenge of poor access to good quality basic seed in order to get certified seed. They then needed a national seed authority (NSA) to certify the seed they multiplied, but the NSA is underfunded and under capacitated to perform this job fully. InovAgro worked with the NSA to design a system for private sector seed inspection under their supervision. Getting the regulations approved by cabinet took two years, but is now operational and is increasing the quantity of certified seed at the national level.

The government of Mozambique did not have the institutional framework needed to respond to the pigeon pea crisis. Even though the trade agreement with India should have allowed Mozambique to continue exporting it did not have the organization to meet the trade agreement conditions. InovAgro and other projects are lobbying for these institutions to be put in place.

A third meso-level challenge is market distortions caused by other donor projects which are pursuing a direct delivery approach. InovAgro has experienced other projects or NGOs working in the same sectors to provide large subsidies to major partners (buying partner participation) or delivering free inputs to farmers to get immediate results (weakening demand from farmers to purchase inputs). This can distort incentives to the market actors over the short term or incentivize SHF to take suboptimal inputs because they are free (also distorting the market for future purchases). While these direct delivery approaches have been proven to not provide lasting results, they also crowd out more sustainable approaches being pursued by InovAgro, causing temporary market distortions that then have to be redressed.

These meso level system changes over time are very important for increased productivity and increased sales to become possible. InovAgro's support to government agencies and lobbying of other donor projects has helped to address some of the distortions.

5. Micro-economic (firm-level) challenges

InovAgro's two earliest partnerships focused on end market buyers seeking to integrate backwards into their supply chains. However, changing corporate strategies in both companies undercut the agreed interventions. One company was bought by a large multinational with a different vision for growth. The other company changed its approach based on the availability of a large grant from another donor to take them in a different direction. A third major lead firm was closed down by its investors in a change in their strategy. This highlights the challenges of finding the right nexus and common agendas between development projects and private sector lead firms, but always being able to adjust to their changes in corporate strategy.

InovAgro's response to these changes in its operating environment was to change its approaches (see below) and focus on seed companies as new leading partners and to develop a new model of output buying.

Adaptation in implementation

The range of challenges above highlights the need for adaptable and opportunistic interventions to address broad systemic constraints to reaching smallholder farmers (SHF). All three of the main interventions where InovAgro has gotten the greatest traction (inputs market, output marketing, and access to finance) required significant adaptation and creativity between the beginning of the impact evaluation and the endline survey.

1. Output marketing

InovAgro's earliest interventions focused on output marketing driven interventions by led by end market buyers, as per above. It became apparent that the large lead firms were not interested in focusing on activities that would target improved relationships with SHF, but rather preferred to

rely on their network of buying agents, made up of Bangladeshis, Chinese, and a few Mozambicans. A new approach was required. The project identified an opportunity through local entrepreneurs who are commodity aggregator traders (CATs) supplying the large end market buyers but were also close to the SHF with an incentive to strengthen their relationships with SHF. This opened up a new channel for reaching backwards to the SHF, increasing transparency and building stronger relations with suppliers.

2. Seed market

After the initial failure of using end market buyers as the key partners to drive productivity increases to SHF, InovAgro seized the opportunity from the government's withdrawal from seed sales to engage the private sector. When first approached in 2011 to determine whether they were interested in marketing directly to SHF, the large seed companies all declined saying they preferred selling in bulk to government. But with the government's withdrawal, many local and international seed companies sought new approaches to reach directly to their end market – the SHF. Starting with two larger seed companies in 2014, InovAgro now works with the eight main seed companies in Mozambique who are driving outreach and training to smallholder farmers. They have a strong incentive to build their market and have proven to be a strong conduit for delivering services to the SHF, organizing 700 demonstrations in 2019 and reaching 18,000 farmers through field days covering good agricultural practices and benefits of certified seed.

3. Access to finance

InovAgro tried working through a range of financial institutions (banks, MFIs, cooperatives) and applying embedded finance through the end market buyers, without getting traction for three years. Ultimately, most financial institutions were not interested in lending to smallholder farmers to help them meet their agricultural input needs without substantial external guarantees, and the regulatory framework (which required IDs and business registrations) were not suitable for SHF to be able to qualify for loans from banks. But finance is required to complete the demand side for improved inputs, so a solution was required.

InovAgro seized on the opportunity presented by the large network of villages savings and loan associations (VSLAs) that were operational in Mozambique to introduce a new methodology focusing on savings for seed – the Fundo Agricola. In addition, the members of the VSLAs were predominantly women, so this was especially empowering for them to access improved inputs. Working through local co-facilitators who are already trainers in the VSLA methodology and knew most of the existing VSLA groups, the Fundo Agricola has met with widespread adoption. With more than 17,400 members (55% of whom were women) now saving for inputs after just four years, the Fundo Agricola is serving as an empowering tool for SHF to both save for agricultural inputs as well as to improve the power balance with large input companies as they negotiate for seed purchases in bulk. It is also serving as the foundation for more last mile retailers to link more closely with their local SHF and to improve their relationships with them.

Lessons learned from implementation

When working through market systems to deliver impact, the behavior of the various actors will drive the results. Therefore, InovAgro had to have a clear value proposition for the partners, both

public and private, which would incentivize them to deliver on the project's desired outcomes of increased productivity and incomes for SHF. Partners need to express the **will** to implement the agreed approaches, as well as the **skill** (technical, financial and managerial capacity) to deliver the results.

A few key lessons stand out from the progress to date.

- Don't rely on just one partner in a sector during the pilot stage; find and support a number of partners to allow for choice, fallback options, and to stimulate healthy competition to drive innovation.
- Establish a clear value proposition for interventions and set out the vision for the future and how the partner would benefit from owning the activities. With a clearly defined value proposition, then define clear roles for the project and partner from the beginning to avoid later confusion.
- Institute continuous monitoring and evaluation of partners and interventions to understand what is actually happening on the interventions and the political economy issues that need to be addressed so that the project can be proactive about fixing them. Understanding the trends will inform the project of changes in value propositions and likely changes in partner behavior.
- Be creative, adaptive, and flexible. If an intervention or a partner is not delivering the desired results, first understand why not and adapt the intervention or the relationship. If that does not work, then be flexible enough to cancel the partnership and move onto trying a new idea. A corollary to this is that, as partner's corporate strategies shift, projects need to understand why they are shifting and be able to pivot with them quickly to continue to influence them on inclusive policies or else pull back support completely.
- Work at the right levels and build the right relationships, especially with government. Much of InovAgro's success has been due to its effective engagement with government agencies (National Director of Agriculture, the NSA, the Ministry of Industry and Commerce (MIC)) who perceive InovAgro to be a partner who is supporting them, not just a source of funding.

VI. Recommendations for future evaluations of MSD programs

1. Recommendations from the evaluator's point of view

- a. Leverage on limited budget and/or administrative capacity of governments, donor agencies and implementing agencies as an opportunity for randomization and implement programs phased-in designs. Better coordination during the planning phase of the project implementation that involves all stakeholders in prioritizing areas of intervention (the donor, implementing firm, and local authorities) may help resolve the ethical concerns in protecting communities as control areas since such strategies can easily be justified as a phased-in process.
- b. Selection bias: Two things may help address potential issues associated with selection bias and contamination issues: (i) Better coordination and commitment from all stakeholders (the donor, implementing firm, impact evaluation team and government authorities) during the planning and implementation phase of MSD programs as well as (ii) the extent to which those managing the experiment are capable and willing to comply with the design of the IE. Not only is designing an incentive compatible situation required (addressing potential conflicts of interest among private sector actors, the impact evaluation team and the implementing firm) but training implementing units on the ground (ie. extension agents and lower level management teams) about the rationale of the design also enhances a better result in the validity of the IE methods.
- c. Utilize proper definition of impact evaluation period as a remedy to account for potential contamination issues or spillover effects. Defining the length of the impact evaluation period depends on the level of impact one seeks to document – i.e., short term/intermediary (such as adoption or behavior change) versus long-term impacts (such as overall income, welfare and productivity impacts). As a rule of thumb, a thorough understanding of the project theory of change should guide the definition of how long the impact evaluation period should be. Hence, impact evaluation period should be long enough to allow the long term impacts to manifest with a bit of caution required since the longer the impact evaluation period, the more susceptible the impact evaluation design will be to challenges of spillover and contamination. If one is only interested in documenting the short term or intermediary outcomes, considering shorter evaluation period is not going to be cost effective but also serve as a methodological remedy to tackle issues of contamination and/or spillover effects.
- d. Adopt a mixed methods approach to account for post-intervention selection bias and contamination issues (for example, randomization, difference-in-difference and propensity score matching techniques used in the case of the InovAgro impact evaluation).
- e. Complementing quantitative impact evaluation methods with a qualitative approach (such as net-mapping of all market actors) could benefit not only in generating methodological remedies to impact evaluation challenges (contamination/spillover effects) but also assess whether the theory of change is

being achieved at macro level (market actors changing their behavior and the system functioning better).

- f. Historical phenomena: Things get complicated if historical phenomena (such as natural disaster) unevenly affect treatment and control groups. In these cases, complementing household survey data with secondary historical data documenting the events occurring during the intervention period may help account for such external phenomena and how these may be contributing to either an under or overestimation of impacts. However, events like currency devaluation are less complicated, as we can expect these effects to be uniform across treatment and control areas.

2. Recommendations from the Implementer's point of view

- a. The implementer must be allowed to present their detailed theory of change to the evaluator and to review the evaluator's target population of interviewees to ensure that they actually reflect the focus of the interventions.
- b. There needs to be a steady interaction with the evaluator to ensure that the evaluator understand the externalities that are occurring which might impact the second and third rounds of surveys.
- c. It is very important to take into account the timeframe for change of behavior and the eventual benefits to the SHF, which can take three or more years – demonstrate seed one year, farmers try it the next year (maybe in the first year if there are giveaway packets for demos), and then expand it the following year. In contrast a project that wants immediate results comes in and gives away seed and training immediately will have impact in the first year. But the impact will not be sustainable beyond the first year and it will most likely distort good behavior needed to put in place the right market systems.
- d. The evaluation needs to have more than just quantitative survey data, but also a lot of qualitative engagement with the market systems actors to see how their behavior is changing.
- e. The results from a market systems project continue well beyond the life of the project. They are long term and sustainable. This can be offset by other, short term behavior by other market actors involved in direct delivery (giveaway) programmes which may yield greater short-term results but which will evaporate when the donor funding ends up. How can the evaluators factor this into the evaluation? It is important for the evaluation to assess the system changes, which will be the long-term drivers of impact on the SHF. This implies that the impact measurement on the on the beneficiaries should be factored to include reviews for after the system has changed, not just while it is changing.
- f. The evaluation needs to be able to incorporate where the PS actually carried out activities that were part of the market systems innovations. Localizing the surveys of the treatment areas before any actual activities are commenced with the partners may distort the results, as implementers do not choose where the private sector actually focus their interventions; while projects can influence the private companies through our agreements, ultimately the private companies will go where

the opportunities are best. In a region that is “virgin” it is critical to get the basic market system established first, which the private sector will concentrate on the higher production zones, before it goes out to the weaker zones. This should go into the selection criteria for the zones to be studied.

- g. Taking into consideration the planned interventions and partnerships, being able to have rolling baselines and evaluations which actually survey farmers where the interventions are occurring is a common practice for project monitoring purposes.
 - a. The populations of the districts are quite large and the exposure of market outlets is limited – there is an average of 300,000 people per district, so if the surveys were carried out in areas not close to intervention initiatives, they would not show results. This can only really be done effectively just before the intervention starts.

VII. References

- Beach, D. and Pedersen, R.B., 2019. *Process-tracing methods: Foundations and guidelines*. Ann Arbor, Michigan: University of Michigan Press.
- Bekkers, H., Miehlabrad, A. and Roggekamp, P., 2008. "How to assess if markets work better for the poor: Experiences from the Katalyst Project in Bangladesh." *Enterprise Development and Microfinance*, 19(2): 120-136.
- Bloom, G., Standing, H., Lucas, H., Bhuiya, A., Oladepo, O. and Peters, D.H., 2011. "Making health markets work better for poor people: the case of informal providers." *Health policy and planning*, 26(suppl_1): i45-i52.
- Brown-Luthango, M., 2010. "Access to land for the urban poor—policy proposals for South African cities." *Urban Forum* 21(2): 123-138.
- de Wildt, M.D.R., 2007. *Accelerating Growth in the Pond Fish Sector: Interventions to bring about sustainable change*. Case Study Number 4. The Springfield Centre for Business in Development, UK.
- Dunn, E., Schiff, H. and Greevey, L., 2011. *Linking small-scale vegetable farmers to supermarkets: Effectiveness assessment of the GMED India project*. Micro Report 166. United States Agency for International Development (USAID).
- Ghebru, H., Smart, J. and Mogue, T. 2019. *Access to markets for smallholder farmers in Alto Molócue and Molumbo, Mozambique: Mid-term impact evaluation of INOVAGRO II*. IFPRI Discussion Paper 1877. Washington, DC: International Food Policy Research Institute (IFPRI). <https://doi.org/10.2499/p15738coll2.133449>
- Gifford, A., DeVries, A., Knott, A. and Mant, H., 2016. *Pioneering New Operating Models and Measurement Techniques for Private Sector-Led Development: Assessing Impact in Nigeria's Niger Delta*. US Initiative for Global Development (IGD).
- Johnson, S., 2013. "From microfinance to inclusive financial markets: the challenge of social regulation." *Oxford Development Studies*, 41(sup1): S35-S52.
- Jones, L. and Oakeley, R., 2012. "Strengthening women's economic empowerment within the M4P framework." *Enterprise Development and Microfinance*, 23(4): 338-352.
- Meyer-Stamer, J., 2006. "Making market systems work? For the poor?" *Small Enterprise Development*, 17(4): 21-32.
- Mills, A., Brugha, R., Hanson, K. and McPake, B., 2002. "Approaches for improving service delivery in the non state sector: what is the evidence on what works, where and why." Paper presented at the World Development Report 2003/4 Workshop held at Enysham Hall, Oxfordshire, UK, November 4-5.
- Oakeley, R., 2011. "Taking stock: Market development-where it is and where it needs to go." *Enterprise Development and Microfinance*, 22(3): 180-183.
- Palmer, N. and Mills, A., 2003. "Classical versus relational approaches to understanding controls on a contract with independent GPs in South Africa." *Health economics*, 12(12): 1005-1020.

Ruffer, T. and Wach, E., 2013. *Review of Making Markets Work for the Poor (M4P) Evaluation Methods and Approaches*. Working Paper 41. UK Department of International Development (DFID).

Tami, A., Mbatia, J., Nathan, R., Mponda, H., Lengeler, C. and Armstrong Schellenberg, J.R., 2005. "Use and misuse of a discount voucher scheme as a subsidy for insecticide-treated nets for malaria control in southern Tanzania." *Health policy and planning*, 21(1): 1-9.

Tschumi, P. and Hagan, H., 2008. *A synthesis of the making markets work for the poor (M4P) approach*. UK Department for International Development (DFID) and Swiss Agency for Development and Cooperation (SDC).

USAID (United States Agency for International Development). 2008. *Cambodia MSME Project Final Monitoring and Evaluation Report*, October.

White, H., 2013. The use of mixed methods in randomized control trials. *New Directions for Evaluation*, 2013(138), pp.61-73.