Juan Bonilla
Nancy McCarthy
Simon Mugatha
Nisha Rai
Andrea Coombes
Joshua Brubaker

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya

May 2018

Impact Evaluation Report 73

Agriculture

3ie 10 YEARS
International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, and use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya*, as partial fulfilment of requirements under grant TW4.2.01 awarded under Agricultural Innovation Thematic Window The content has been copy-edited and formatted for publication by 3ie. Despite best efforts in working with the authors, some figures and maps could not be improved. We have copy-edited the content to the extent possible.

The 3ie technical quality assurance team for this report comprises Diana Lopez-Avila, Deeksha Ahuja, Stuti Tripathi, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Emmanuel Jimenez The 3ie editorial production team for this report comprises Angel Kharya and Akarsh Gupta, with Beryl Leach providing overall editorial supervision.

All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Please direct any comments or queries to the corresponding author, Juan Bonilla at jbonilla@air.org.

Funding for this impact evaluation was provided by International Fund for Agricultural Development.

3ie's Agricultural Innovation Thematic Window is supported by the Alliance for Green Revolution in Africa, the Bill and Melinda Gates Foundation, the International Fund for Agricultural Development and UK aid through the Department for International Development.

Suggested citation: Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018. *Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73.* New Delhi: International Initiative for Impact Evaluation (3ie). Available at: doi:https://doi.org/10.23846/tw4IE73

Cover photo: Riccardo Gangale / USAID

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya

Juan Bonilla
American Institutes for Research (AIR)

Nancy McCarthy
LEAD Analytics

Simon Mugatha
AIR

Nisha Rai
AIR

Andrea Coombes
AIR

Joshua Brubaker
LEAD Analytics

3ie Impact Evaluation Report 73

May 2018

Acknowledgements

We recognise the contributions of many individuals and organisations, without whom it would not have been possible to complete this study. Our thanks go to the International Initiative for Impact Evaluation (3ie), the Smallholder Dairy Commercialization Programme (SDCP), the International Fund for Agricultural Development (IFAD), the International Livestock Research Institute (ILRI) and Research Solutions Africa (RSA) for the opportunity to carry out this study, and the financial and technical support that they provided. We further appreciate 3ie for the technical review and support provided throughout the study.

Our special thanks go to Jasper Grosskurth and Jasper Gosselt (RSA) for carrying out the data collection, and to Pamela Ochungo for her help with re-creating the original ILRI targeting maps. The value of the technical and logistical support obtained from Luke Kessei cannot be overemphasised. Everyone at the SDCP provided valuable logistical support during data collection, including programme background information.

Our acknowledgments would be incomplete without mentioning our team of very able research assistants in Kenya. Specifically, we acknowledge the input of the team of enumerators and supervisors from RSA, whose dedication during data collection ensured that high-quality data were collected. We also gratefully acknowledge the highly competent team of programming and data cleaning personnel at RSA.

Finally, we wish to acknowledge the patience that the Kenyan households, community leaders and community members exercised during interviews. It is our hope that insights from the information they provided will translate into valuable interventions in their communities.

Summary

Programme overview

The dairy sector in Kenya is one of the largest and most developed in Sub-Saharan Africa, accounting for 4 per cent of its gross domestic product. Despite high production volumes, the sector is still dominated by smallholder farmers who rely on livestock for income and food security. Dairy farmers face a number of barriers to increased profitability, including: animal diseases; lack of access to artificial insemination and other veterinary services; high costs of improved technologies, such as silage equipment; inadequate access to markets; poor rural infrastructure; and unsteady supply of quality animal fodder.

The Smallholder Dairy Commercialization Programme (SDCP) was funded by the International Fund for Agricultural Development (IFAD) and implemented by the Government of Kenya from 2005 to 2015. It was designed to reach 600 dairy groups (24,000 smallholder dairy farmers) in nine counties. SDCP provided training to dairy farmers to build their enterprise, managerial and organisational skills. Aside from training, the programme also aimed to enhance dairy farming productivity and reduce production costs through demonstrations, field days and grants. To strengthen market linkages, SDCP invested in improving road infrastructure and conducted additional training on milk-handling practices and value-added opportunities.

The programme identified three main areas where barriers to improving dairy income potentially occur: dairy group activities, household production and market intermediaries. Programme designers hypothesised that increasing net dairy income for smallholder farmers can occur through four primary contextual factors: (1) increasing milk production; (2) increasing milk prices; (3) decreasing the costs of producing milk; and (4) decreasing the transaction costs of participation in input and output markets. They assumed that increased net income would lead to improved food security and increased participation by women and marginalised communities.

Impact evaluation overview

A key objective of agricultural extension is to increase farmers' knowledge about agricultural practices, which in turn could have an effect on productivity. Evaluating the impact of the SDCP can provide useful insights for the design of other agricultural extension programmes for smallholder farmers in developing countries. 3ie funded this grant under its Agricultural Innovation Thematic Window. The main evaluation questions were the following:

1. Does SDCP improve the wellbeing of dairy farmers through improved animal management, improved efficiency, increased production and productivity, and an increase in farmers' incomes and food security?
2. Are there differences in the participation in programme activities between female and male dairy farmers? Does the programme have sex-differentiated effects?
3. How does the process by which SDCP is implemented influence the effectiveness of the programme? Given this process, how do contextual factors affect programme success?

Methodology and identification strategy

The ex-post evaluation used a quasi-experimental design using propensity score matching. The sample of 2,500 famers was split equally between 1,250 SDCP beneficiaries and 1,250 non-programme participants. Key information was collected through a survey (conducted by Research Solutions Africa), and the programme causal chain – including underlying assumptions – was mapped out. To capture key time-invariant characteristics, as well as retrospective information on dairy farming, the study included a short filter questionnaire at the beginning of the household survey for accurately matching households in the treatment group with households in the comparison group.

The study used qualitative data to explain the quantitative findings and understand the implementation of the programme. It compared accounts of local government officials with those of dairy farmers by using semi-structured interviews and focus group discussions.

Main findings

The findings suggest that the SDCP was successful in increasing milk production, but had limited impacts on increasing either the quantity of milk sold or prices received by famers. Regarding intra-household bargaining power, SDCP farmer households were more likely to have women managing cash from the sale of milk, relative to non-SDCP households. This was also found to be true for decisions relating to the use of services such as artificial insemination, anthelmintic drugs, tick control, vaccination and curative treatments.

Qualitative research highlighted the challenges in programme implementation, such as those related to linking dairy groups to various service providers, and limited knowledge among farmers to negotiate terms favourable to them.

We calculate that it would take approximately 4.74 years for the programme to break even; that is, for the benefits to equal the costs, assuming the benefits of increased milk production remain constant across years. This estimate of the number of years to break even seems reasonable, for although it was unlikely there were benefits at the beginning of the programme (when upfront costs were associated with setting up the SDCP), farmers were likely to have started benefiting from the programme before 2016. Furthermore, to the extent that farmers continue to employ the best practices the programme advocates, the benefits may extend into future years.

Contents

List of figures and tables

Acronyms

3ie	International Initiative for Impact Evaluation
AI	Artificial Insemination
AIR	American Institutes for Research
DCA	Dairy Commercialization Area
DFID	Department for International Development
FANTA	Food and Nutrition Technical Assistance Project
FGD	Focus Group Discussion
GDP	Gross Domestic Product
GoK	Government of Kenya
IFAD	International Fund for Agricultural Development
ILRI	International Livestock Research Institute
IPWRA	Inverse-Probability-Weighted Regression Adjustment
IRB	Institutional Review Board
RSA	Research Solutions Africa
SDCP	Smallholder Dairy Commercialization Programme
T&V	Training and Visit

1. Introduction

The International Initiative for Impact Evaluation (3ie), on behalf of the International Fund for Agricultural Development (IFAD), contracted the American Institutes for Research (AIR) in partnership with LEAD Analytics and Research Solutions Africa (RSA) to evaluate the Smallholder Dairy Commercialization Programme (SDCP) in Kenya.

The SDCP, implemented by the Government of Kenya (GoK) from 2005 to 2015, was designed to reach dairy groups engaged in milk production in nine milk-producing counties. The overall goal of SDCP was to increase the income of poor rural households that depend substantially on production and trade of dairy products for their livelihoods. The SDCP was implemented through various interrelated components.

First, SDCP provided training on organisational, managerial, and enterprise skills (e.g. bookkeeping, accounting and financial planning) to farmers. Second, the programme targeted household production by aiming to enhance dairy farming productivity and reduce production costs through grants, trainings, field days and demonstrations. Last, SDCP aimed to strengthen market linkages for small-scale milk producers by improving road infrastructure and conducting additional trainings on milk handling practices and value-addition opportunities.

Many researchers, non-governmental organisations, governments and donors have long held the position that smallholder dairy can be a particularly effective mechanism for alleviating poverty and increasing food security in regions well-suited for dairy production, such as those located in western Kenya (Staal et al. 1997; Thorpe et al. 2000; Burke et al. 2007). While there had been some success before 2006 in promoting the production and marketing of milk in these types of regions in Kenya, the SDCP was designed to address a number of barriers to increased profitability of smallholder dairy farming in the country. These included relatively high transaction costs for production and marketing (Staal et al. 1997), and the underperformance of dairy groups and cooperatives in reducing those transaction costs (Holloway et al. 2000; Atieno and Kanyinga 2008).

However, empirical evidence on the impacts of reducing transaction costs – either through analyses of household surveys or through rigorous impact assessments of specific projects – remains relatively limited. A few studies report results that would suggest that reducing transaction costs improves farm productivity and incomes, primarily by increasing access to improved production methods and marketing information, and by being able to link to private sector actors in the value chain.

Burke et al. (2015) find that smallholders closer to electricity sources and with access to private sector value chain actors are more likely to be dairy producers and net sellers, and that active dairy cooperatives also induce greater participation by smallholders in dairy markets. Gelan and Muriithi (2015), and references cited therein, find that practices such as zero grazing and the adoption of improved cows led to increased milk production efficiency, while Nafula (2013) provides evidence of the importance of enhanced nutritional and feed practices on the health of dairy calves, using a randomised control trial method. The SDCP project promotes all of these activities and also seeks to reduce marketing-related transaction costs.

Understanding the impacts of the SDCP is particularly valuable, given the importance of the dairy sector in Kenya. The dairy sector in Kenya is one of the largest and most developed in Sub-Saharan Africa, accounting for 4 per cent of the country's Gross Domestic Product (GDP). Despite high production volumes, the dairy industry is dominated by smallholder farmers who rely on livestock for income and food security.

However, smallholder dairy farmers face a number of barriers to increased profitability, including: irregular supplies of good-quality animal fodder and feed throughout the year; animal diseases; inadequate access to markets; lack of access to Artificial Insemination (AI) services and other veterinary services; poor rural infrastructure; high costs of improved technologies, such as silage equipment; limited skills with which to bargain with input suppliers and output purchasers; limited knowledge about how to run their dairy activities as commercially oriented enterprises, including maintaining animal health and providing high-quality milk; dependency on traders; and limited ability to maintain milk quality, putting downward pressure on farm gate milk prices.

While the project undertook a study of the programme in the form of household surveys at the end of the project in 2015, no formal impact assessment had previously been conducted. Specifically, the 2015 survey did not include households from 'control' areas. Instead, the questionnaire included retrospective questions, going back many years. Thus, it is possible that findings from the 2015 survey might suffer from recall bias and from lack of data from households on which to construct a counterfactual.

The study AIR and LEAD Analytics have conducted is the first rigorous impact assessment designed to examine the effects of the SDCP. AIR and LEAD Analytics designed the evaluation to address the knowledge gaps that related to the project's impacts on smallholder dairy producers, including measures of production efficiency and profitability, as well as impacts of project activities aimed at ensuring that women and resource-poor farmers benefitted from the project.

The study investigated programme impacts and implementation of the SDCP. Our main research question was whether SDCP improves the wellbeing of dairy farmers through improved animal management, increased production and productivity, improved efficiency (e.g. input and transaction cost reductions), and a rise in farmer incomes and food security. We also investigated whether there are differences in programme participation and effects by the sex of the farmer. Lastly, we looked at the process by which SDCP is implemented and how contextual factors may affect programme success.

We addressed these research questions through a mixed methods design. Quantitatively, we used a quasi-experimental approach using a matching ex-post design. Since programme implementation is now complete, an ex-post analysis was appropriate to inform future scale-up of similar efforts. Our design involved two steps of matching at division and then household level. At the division level, we relied heavily on a two-step targeting approach.

First, we used existing administrative information from the original programme targeting to identify similar non-programme areas that were not affected by the SDCP due to capacity constraints (i.e. observation-based targeting). Second, we discussed the filtered control communities with local experts and stakeholders to determine which of the pre-

selected control areas were more similar to treatment communities at project inception (i.e. criteria-based targeting).

One of the major benefits of the criteria-based targeting exercise has been that, based on the experts' insights, we have been able to exclude divisions where other policies or actions may have had differential effects over time, which has been especially beneficial since this is an ex-post evaluation for a long-running programme. Furthermore, we matched treatment and comparison divisions within the same counties, which has helped ensure that both study areas receive the same level of support from their county over time, especially after devolution of power to the counties, which occurred in Kenya in 2013.

At the household level, to estimate programme impacts, we further used matching through our use of the doubly robust estimator, which combines regression and propensity score methods to achieve some robustness to misspecification of the parametric models (Imbens and Wooldridge 2009). Combining regression and weighting (with weights derived from the probability of being part of the programme) can lead to additional robustness by removing the correlation between omitted covariates, and by reducing the correlation between omitted and included variables. We complemented these analyses with qualitative research in the form of Focus Group Discussions (FGDs) and key informant interviews.

Our findings indicated that the programme had statistically significant positive impacts on improved animal management, including grazing and feeding practices, keeping practices and animal health services. These results suggest that farmers were receptive to programme education regarding the nutritional content and feeding practices, which may have a positive effect on milk production. The SDCP also led to an improvement in breeding services. Specifically, the SDCP increased the propensity of households to monitor their cattle on a regular basis and use AI services. The positive quantitative findings related to access to and use of AI are especially promising, considering that many farmers discussed ongoing challenges with AI.

We found that SDCP treatment farmers have improved access to extension visits, field days and demonstrations – services that the programme provided. We also found an increase in the probability of receiving any information on specific aspects of the production process. The largest increase in probabilities was for receiving information on general livestock practices, milk processing and quality control, fodder establishment and fresh milk marketing. However, we observe positive impacts for almost all types of cattle-related topics, as well as topics related to enterprise skill. The results also suggest that the SDCP was successful in increasing milk production, but with more limited – though positive – impacts on increasing milk marketing and increasing milk prices received by smallholders.

The report is structured as follows. First, we describe the intervention and research hypotheses. Then we describe the context and the timeline for the project and evaluation. We describe the design, methods and implementation of the evaluation and policy. We then present the impact analysis, followed by a discussion and recommendations for policy and practice.

2. Intervention, theory of change, and research hypotheses

2.1 Development of intervention

During the 1990s, the monopolistic dairy cooperative structure in Kenya collapsed. While plagued by inefficiencies, the structure provided some support to dairy farmers. The collapse led to great distrust of new collectives, despite the potential gains from collective action in the sector and due to economies of scale in linking smallholders to input suppliers and milk purchases (e.g. through bulk purchasing of fodder and bulk milk sales).

At the outset of the project in 2006, the GoK was restructuring at least some of its legislation related to smallholder participation in dairy markets, with the aim of reducing costs that smallholders face in joining formal markets. One of the project's goals was to work with the government and relevant ministries to further enhance the policy and legislative environments for smallholders, primarily with respect to issues dealing with animal breeding.

The project was built on empirical evidence on smallholder dairy systems the International Livestock Research Institute (ILRI) produced, with support from Department for International Development (DFID). Given evidence on the level of dairy commercialisation's potential, smallholder dairy producers, poverty rates and indicators of rural infrastructure, the project chose to work in nine districts in central and central-western Kenya. Smallholders dominate the production of milk in these districts, which also exhibit high poverty rates.

At the same time, smallholder dairy farmers face a number of barriers to increased profitability, including: irregular supplies of good-quality animal fodder and feed throughout the year; animal diseases; inadequate access to markets; lack of access to AI services and other veterinary services; poor rural infrastructure; high costs of improved technologies, such as silage equipment; limited skills with which to bargain with input suppliers and output purchasers; limited knowledge about how to run their dairy activities as commercially oriented enterprises, including maintaining animal health and providing high-quality milk; and depending on traders with limited ability to maintain milk quality, putting downward pressure on farm gate milk prices.

Additionally, women play a key role in the smallholder dairy sector. Research showed that many female-headed households in these districts had dairy cows, and even in male-headed households females controlled over 60 per cent of the income from dairy activities. Women dairy farmers traditionally have been even further disadvantaged in terms of receiving extension advice and playing leadership roles in dairy groups to ensure their specific needs are addressed.

The SDCP – funded by IFAD, the GoK and the local community, and implemented by the GoK from 2005 to 2015 – was designed to address some of the constraints smallholder dairy farmers faced. The overall goal of the SDCP was to increase the income of poor rural households that depend substantially on production and trade of dairy products for their livelihoods.

The SDCP was implemented through various interrelated components. First, the programme provided dairy group activities by training beneficiaries on organisational, managerial and enterprise skills (e.g. bookkeeping, accounting and financial planning) to fully benefit from market-driven milk commercialisation. Further, capacity building of dairy groups was complemented by competitive access to investment grants for improving forage production and feed milling, milk bulking, chilling and processing, and management and market information systems.

Second, the programme targeted household production by aiming to enhance dairy farming productivity and to reduce production costs through trainings, field days and demonstrations.

Third, the SDCP aimed to strengthen relationships with market intermediaries by enhancing market linkages for small-scale milk producers, improving road infrastructure, and conducting additional trainings for beneficiaries on milk handling practices and value-addition opportunities.

The primary beneficiaries of the project were resource-poor smallholder dairy farmers, with an emphasis on ensuring women's participation in all project activities. The project intended to reach 24,000 smallholder dairy-farming households, with members participating in 600 dairy groups across 9 milk-producing counties. IFAD and the GoK determined the target of 600 dairy groups as a number large enough to have an impact, but still within capacity constraints.

Based on dairy group inventories as of 2011 and 2012, the project ended up working with 15,535 smallholder dairy-farming households, with members from 505 dairy groups across the 9 counties. In addition to resource-poor smallholders (85 per cent of targeted beneficiaries), the project also targeted smallholders already engaging in more intensive production.

Finally, the project also targeted 300 milk traders and 90 milk bar/milk processors to improve milk quality and contractual arrangements with buyers and sellers.

2.2 Theory of change

The project first reviewed the substantial evidence on factors associated with low milk production, productivity per cow, and relatively low participation in milk markets with marked seasonal fluctuations. The project identified three main areas where impediments to improving dairy incomes occurred: dairy group activities, household production and market intermediaries. These three areas conform to Components 1–3 of the five project components, with support to policy and institutions and to project management comprising the fourth and fifth components.

Increasing dairy incomes for smallholders can occur through three primary channels: increasing milk production, increasing prices received for milk sold, and decreasing costs of producing and marketing milk. In this evaluation, we focus on the first three components, as these address all of the channels to varying degrees.

Component 4 (support to policy and institutions) concerns activities to shape national-level policies and regulations, primarily concerned with regulating AI, registering

(improved) breed births, and milk-related phytosanitary regulations. To the extent that these have been adopted and implemented, they would affect all dairy cattle owners in Kenya; that is, dairy cattle owners in both the treatment and comparison groups in this study. Since both groups would benefit from these policy and regulatory changes, it is impossible to evaluate these impacts using the treatment and comparison households. Component 5 (support to project management) affects only treated households, so it is more of an IFAD-specific procedure than a separate activity. In fact, most development projects do not put project management as a separate component of activities.

In other words, all beneficiaries receive all five components, whereas comparison farmers would receive benefits from Component 4. Since beneficiaries receive all five components, isolating their effects is not feasible with the quantitative evaluation. However, since the main difference between treatment and comparison groups is found in Components 1–3, we focus on these components. These components are linked, because farmer-level interventions (Component 2) interact with how groups of farmers perform (Component 1) in order to be able to take advantage of greater market opportunities (Component 3).

2.2.1 Component 1: Organisation and enterprise skills of dairy groups

Main barriers:

a. Disorganised groups with limited enterprise/commercialisation skills and knowledge;
b. Limited or weak links with input suppliers and output purchasers, leading to missed opportunities to secure lower input prices or higher and more reliable output prices; and
c. Limited ability to disseminate relevant production advice to farmers, particularly disadvantaged groups including women and resource-poor farmers.

Main inputs to address these three barriers:

I1. Extensive training on group organisation and management, enterprise skills, development of enterprise plans, and preparation of business proposals that are then eligible for dairy enterprise grants;
I2. Training on establishing and maintaining links with input and service providers and output purchasers;
I3. Linking groups to advisory and extension systems; and
I4. Reaching out to women and resource-poor farmers.

Main intended outcomes:

O1. Dairy groups with financially viable and sustainable business plans, and the ability to develop and successfully obtain external grant funds (due mainly to I1 input). The main assumption is that dairy group members were able to successfully understand the training materials and translate that knowledge into business plans and proposal writing. Another key assumption is that there were real business opportunities that relatively small and resource-poor dairy groups could take advantage of;
O2. Transaction costs and input costs reduced, and output prices potentially increased (due mainly to I2, and to some extent I1). The main assumption is that access to milk markets and market players can be increased by knowledge gained in training. It should be noted that the project was undertaken in areas

determined to be high-potential commercialisation areas, which was intended to limit the impact of other barriers such as geographic isolation and very high transportation costs; and

O3. Dairy group members, including women and resource-poor farmers, increase knowledge on dairy production and markets, leading to higher and more stable milk production and to lower transaction costs of participating in markets (due mainly to I3, and to some extent I1).

2.2.2 Component 2: Technical support to dairy producers

Main barriers:

a. Poor breeds of dairy cows;
b. Lack of adoption of improved management practices such as zero grazing, keeping milk records, and producing hay and silage;
c. Poor animal nutrition, and pronounced seasonality in fodder and feed use; and
d. Animal diseases.

Main inputs to address these four barriers:

I1. Trainings and demonstrations to disseminate information about benefits to improving breeds through AI, benefits from animal registration, improved husbandry and dairy enterprise management practices, improved fodder production and management, supplemental feed use, and animal diseases and disease management;
I2. Establishing community AI schemes;
I3. Establishing a revolving community-based animal health fund; and
I4. Training community resource persons to aid in disseminating information and linking farmers to relevant resources.

Main intended outcomes:

O1. Better-bred dairy cows, leading to higher milk per cow and total output (I1 and I2). The main assumption is that quality AI seed is available, and that farmers see the value in improving breeds, which provides delayed benefits;
O2. Greater production and better management of fodder and feed, leading to lower costs of milk production and greater stability in milk output throughout the year (I1 and I4). The main assumption is that limited opportunity costs associated with fodder being put to other uses and options for extending forage availability throughout the year (e.g. storage) are profitable;
O3. Healthier cows producing more milk (I1, I3 and I4). The main assumption is that information on disease management and access to revolving funds are sufficient to address substantial issues with tick-borne disease control;
O4. Better overall management practices, leading to greater production and potentially lower costs of production (I1 and I4). The main assumption is that training materials contain relevant and understandable information that farmers can apply in practice. It also assumes relatively low opportunity costs, investment costs and input costs, or alternatively, that switching to new systems is actually profitable; and
O5. More milk to meet household needs and to participate year-round in the milk market (I1 enterprise management and fodder production and management, and to some extent I4). The main assumption is that market linkages are established (e.g. through the dairy groups) and transaction costs are sufficiently lowered.

2.2.3 Component 3: Development of milk marketing chains

Main barriers:

a. Disorganised markets;
b. Market traders with limited skills to maintain high-quality milk (e.g. hygiene standards);
c. Market traders with limited enterprise management skills;
d. Limited generation and dissemination of milk market information;
e. Limited market infrastructure (e.g. bulk milk facilities);
f. Limited production and marketing of dairy goat milk;
g. Inefficient contractual arrangements between dairy groups, milk collectors, cooling centres and processors; and
h. Limited access for smallholders to rural finance.

Main inputs to address these eight barriers:

I1. Developing a low-cost market information system and strengthening the dairy information centre;
I2. Linking activities between smallholders and rural finance operators;
I3. Capacity building for milk marketing groups;
I4. Pilot testing school milk programmes as an opportunity to expand milk marketing;
I5. Performing a study on milk marketing opportunities;
I6. Trainings on hygienic milk handling;
I7. Establishing milk bulking facilities and other infrastructure (e.g. cooling facilities);
I8. Training and demonstrations on dairy goat production and marketing, and procurement and distribution of dairy goats to resource-poor smallholders; and
I9. Developing new and improved contractual arrangements

Main intended outcomes:

O1. Reduced transaction costs of participating in the market (I1, I2, and I7). The main assumption is that farmers and other market players can access new information sources, and that information is relevant and understandable;
O2. Increased size of the market (I3, I4, and I5). The main assumption is that there is scope for expanding the milk market. External evidence suggests there is such scope;
O3. Increased quality of milk in the market, increasing the value added and potentially leading to higher prices for smallholders (I6 and I7). The main assumption is that knowledge is disseminated in a practical and useful way, and more importantly, that sufficient access to technologies and infrastructure exist throughout the entire value chain;
O4. More effective contractual arrangements to increase the quantity of milk in the market throughout the year. The main assumption is that contract terms are currently inefficient and there is scope to make improvements; and
O5. Increased participation in the dairy goat milk market by women and resource-poor smallholders (I8).

2.2.4 From outcomes to impacts

As noted above, the primary impact is expected to be higher net milk incomes through increased production and productivity per animal, reduced input costs, reduced transaction costs, and potentially higher farm gate milk prices. The outcomes clearly map to this impact. The second expected impact is greater participation by women and

resource-poor farmers in milk markets and as leaders in dairy groups, which follows from Component 1, Outcome 3 and from Component 3, Outcome 2. While we can analyse this impact through heterogeneous effects analysis, the greater participation by women and resource-poor farmers is an impact in and of itself, since the project aims to ensure that women and the most vulnerable are indeed included. The third expected impact is increased food security. In part, this is related to higher net dairy incomes, and thus related to two of the four pillars of food security: access and availability. Additionally, greater stability of milk production and sales throughout the year is related to the stability pillar. Finally, trainings on hygiene increase milk safety for consumers, which is related to the utilisation pillar.

2.3 Primary outcomes and impacts of interest

The programme's key intended outcomes included more knowledgeable dairy farmers; healthier and better-bred milk cows; reduced seasonality in milk production; increased participation and labour opportunities in milk markets (generating the impact of higher net dairy incomes for smallholder dairy farmers); well-developed business plans and project proposals; establishment of reliable trade relations with input suppliers and output purchasers for dairy groups (generating the impact of commercially viable and sustainable dairy organisations); and increased networks and higher quality milk for milk traders (generating the impact of greater value addition).

In this study we investigate programme impacts and implementation of the SDCP. Our main research question is the following:

1. Does the SDCP improve the wellbeing of dairy farmers through improved animal management (e.g. better husbandry practices), increased production and productivity (e.g. litres/cow/day), improved efficiency (e.g. input and transaction cost reductions), and increased incomes for farmers (e.g. gross margins and higher milk prices)? Does this lead to increased income from dairy farming and ultimately improved food security?

To better understand our primary research question, we consider the following secondary questions:

1. Are there differences in programme participation by the sex of the farmers? Does the programme have differential effects by the sex of the farmers?
2. How does the process by which the SDCP is implemented influence the effectiveness of the programme? Given this process, how do contextual factors affect programme success?

The first secondary question is critical, since women play a key role in dairy production in Kenya. However, they face a number of constraints that may alter the extent to which they benefit from the programme, such as owning smaller farms, which affects their access to credit using land as collateral, or being less educated, which limits their access to technical information for enhancing production.

The second secondary question is especially important, because the impacts of a programme are ultimately a function of the manner in which it is implemented. Understanding the implementation of this programme is a critical aspect of this evaluation because it largely determines the type of dairy farmer the programme

reaches. There are clear implications if the programme does not reach certain segments of the population (e.g. women, poorer farmers, certain locations, etc.). The process can also lead to variation in programme effectiveness in that it may vary according to contextual factors, such as the population density of a region or the type of f arming.

3. Context

Agriculture is one of the leading sectors in Kenya, employing 70 per cent of the rural population and accounting for 25 per cent of GDP. More specifically, the dairy sector in Kenya is one of the largest and most developed in Sub-Saharan Africa, accounting for 4 per cent of the country's GDP. Dairy farms are concentrated in the highlands and former provinces of the Rift Valley in the central and eastern parts of Kenya. There are more than 3.5 million head of purebred Friesian-Holstein, Ayrshire, Guernsey and Jersey cattle and their crosses, with a total yearly production of about 2 billion litres of milk (Muriuki 2003). Despite these high production volumes, smallholder farmers dominate the dairy industry.

Approximately 80 per cent of people engaged in agricultural activities are smallholder farmers, producing three-quarters of the agricultural output in farms that commonly have an area of less than 3 ha. Smallholders produce over 50 per cent of the predominant crops in the country – including maize, coffee and tea – and 80 per cent of milk and beef products. There are over 1 million smallholder farmers who depend on dairy farming for their livelihoods. Less than 15 per cent of marketed milk flows through milk processors (Thorpe et al. 2000), with the rest being sold as raw milk through direct sales to consumers by farm households, dairy cooperative societies and individual traders.

As with many other countries in the Sub-Saharan region, smallholder dairy farmers in Kenya rely on livestock for income and food security. However, milk production is regularly threatened by inadequate access to markets; animal diseases; poor quality and unstable supplies of animal food and feeds, with corresponding decreases in market milk supply; poor rural infrastructure; and inadequate access to AI services and other veterinary services.

More importantly, small dairy farmers face a series of constraints that prevent them from effectively commercialising their milk, including large seasonal fluctuations in milk output and prices; poor rural infrastructure (e.g. roads and electricity); limited skills with which to bargain with input suppliers and output purchasers; lack of management and enterprise skills; and inefficiencies in the post-harvest segment of the dairy value chain.

Furthermore, the dependency on traders who have limited ability to maintain milk quality puts downward pressure on farm gate milk prices. The inability of farmers to avoid and respond to these risks and constraints associated with milk production is partly explained by the lack of access to productive inputs and improved technologies, such as silage equipment, and also to a lack of knowledge on how to maintain animal health and increase milk productivity through better production practices.

3.1 Programme targeting

In 2005, IFAD and the GoK commissioned a study by ILRI to lead the initial targeting of the programme. Three key indicators were initially considered to determine the

programme area: milk production and production potential, incidence of rural poverty, and market access. Milk production was measured as litres of milk produced per square kilometre per year – an indicator that reflects well the productivity by individual cow, number of animals in a given area, and percentage of lactating animals. Poverty incidence was drawn from data constructed by the Central Bureau of Statistics in 2003. Finally, market access was defined as distance from farms to main milk cooling centres, as well as distance from farms to main urban centres.

On realising that the three key indicators were too restrictive in terms of potential programme locations, the programme targeting instead focused on the two key indicators of milk production and production potential, and incidence of rural poverty. For these indicators, thresholds were defined to target the initial participating districts/counties. First, the programme selected districts with high (greater than 90,000 litres/km^2/year) and medium (less than 90,000 litres/km^2/year) milk densities. Second, they chose districts with a poverty rate of at least 46 per cent.

To select beneficiaries, the project relied on 'Targeting Pro-Poor Investment in the Dairy Sub-Sector', a report ILRI produced in March 2005, which contained rich analysis of the smallholder dairy subsector. Using additional information from dairy farmer surveys, the project first selected 9 districts out of the 25 milk-producing districts in Kenya. Then, because not all divisions/sub-counties (the next administrative unit after districts/counties) within the 9 selected districts met the targeting criteria, the SDCP focused operations on only 27 divisions (out of a total of 53 divisions in the 9 districts). These divisions corresponded to Dairy Commercialization Areas (DCAs), and the GoK selected locations and sub-locations within the 27 DCAs, with the aim of identifying 600 farmer groups. The project attempted to work with existing farmer groups that had at least some activities focused on dairy farming. Each DCA had 500–800 dairy farmers, which was considered a manageable number of farmers for divisional government staff to handle.

Once the geographical units for the programme were selected, the SDCP put programme activities into operation in the field through dairy groups. Dairy groups are formally registered as self-help groups, cooperative societies or common interest groups that have a common interest in dairy farming. On average, each SDCP dairy group has 30 dairy farmers. By the end of 2015, the SDCP had worked with 527 dairy groups, with more than 16,000 members, of which 60 per cent were women.

Selection of dairy groups within programme areas was done using a participatory selection process that gave priority to resource-poor dairy farmers, including women and youth. More specifically, the SDCP chose individual dairy groups based on the following observable characteristics: (a) farmers had a maximum of two cows; (b) farms had an average of 2.5 acres (0.91 ha) of productive land; and (c) the group was not marketing more than 30 per cent of its milk before joining the programme. The programme targeted the dairy group; thus, all participating farmers in the dairy group were eligible.

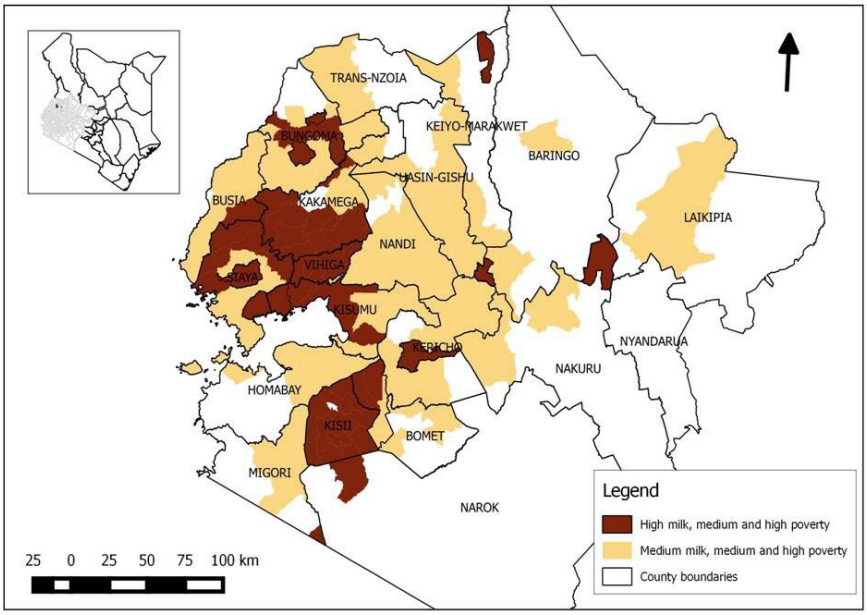
Lastly, the identification of existing dairy groups in each district was facilitated by the fact that, in Kenya, all common interest groups need to be registered with the national and local governments. While IFAD and the GoK targeted 600 dairy groups, numerous other dairy groups existed in the districts that were not treated due to limited implementation.

3.2 Study site selection

To select the sites for the study, we replicated the targeting process that ILRI conducted in 2005. This replication process involved the use of observation- and criteria-based targeting, as proposed by Ouma and colleagues (2007). Observation-based targeting involved determining where the SDCP was adopted, plotting those sites on a map, and identifying the common characteristics the sites shared.

To conduct the observation-based targeting, the AIR and LEAD Analytics team worked with contacts at ILRI to examine the data that were originally used for the targeting exercise. We engaged with Pamela Ochungo, a geographic information systems analyst from Kenya who conducted the original targeting exercise on behalf of ILRI, to recreate the maps used for the original targeting of the programme. Figure 1 shows the areas that met the threshold for high and medium milk production and poverty levels.

Figure 1: Recreation of original programme targeting



We used this information to determine a group of potential non-programme areas that had characteristics similar to SDCP areas before the programme started. Because IFAD and the GoK determined the target of 600 dairy groups in part due to capacity constraints, there were additional non-treated areas similar to the treated areas that could serve as a comparison group. After recreating the original targeting map, we performed propensity score matching at the level of the division by using the original data to match treatment divisions to potential comparison divisions within each county.

We calculated from the original ILRI data a propensity score for the division, based on the 2005 values of the variables of milk density, proportion of poor households, proportion of grade cows in dairy households, proportion of dairy households, and travel distances to urban and cooling centres. We matched similar comparison divisions to treatment divisions within the same county; that is, we simultaneously selected treatment and comparison divisions for the study areas on the basis of their being the most comparable pairs according to observed 2005 data. Restricting our comparison to the same county was important, because over the programme implementation years county-

level support for dairy farming varied after devolution. Thus, to mitigate specific influences at the county level, we needed to compare treatment divisions with comparison divisions in the same county.

In addition to the results of the matching exercise, we used criteria-based targeting to refine the selection of the comparison sites. Criteria-based targeting is based on the opinion of experts, who determine to what extent non-targeted areas could have been chosen for the programme. Experts relied on historic variables that are likely to be associated with the uptake of the intervention in 2005, such as climate, market access, and other agro-ecological conditions.

Luke Kessei, the ministry desk officer for the SDCP from 2005 to 2015, played a key role in designing and implementing the programme and is now an independent consultant. He, along with the SDCP technical team, served as our experts for the criteria-based targeting. In October 2016, we met with IFAD and SDCP officers in Nairobi and Nakuru. Through the meetings with the SDCP technical team, we discussed at a high level which of the potential comparison areas might serve as the best counterfactual in terms of similarities to SDCP treatment divisions.

We presented to the team a list of the nine programme counties along with the potential divisions (all the beneficiary divisions that could serve as study treatment areas and the non-beneficiary divisions that could serve as comparison areas). With the help of our experts, we narrowed down the list of SDCP treatment and comparison divisions to include in the evaluation. In these discussions, the experts considered eight criteria:

1. Whether the evaluation included at least one county from each of the three geographic programme clusters;[1]
2. Variation in county-level support after devolution, where areas with high levels of support were excluded;
3. Variation in support provided by other dairy-focused programmes, where areas with high levels of support were excluded;
4. The violence that happened in certain areas during and after the 2007 elections, which resulted in factories being destroyed, livestock being killed, and some areas receiving government support in response to the violence. Areas that had very high levels of violence were excluded;
5. Similarities in farmer composition, since some potential comparison divisions had large-scale farmers that would not serve as a valid counterfactual to the smallholder farmers participating in the SDCP. Areas where the farms tended to be larger in scale were excluded;
6. Geographic proximity of the potential comparison areas to the SDCP treatment areas, since comparison areas that neighboured the SDCP areas would be at higher risk of contamination from spillover effects. Comparison areas that closely neighboured SDCP treatment areas were excluded;
7. The focus of the farmers in the area, since some farmers were focused more on tea farming than on dairy farming. Areas where the farmers tended to be mostly tea farmers were excluded; and
8. The feasibility of data collection requirements in very remote divisions, which were excluded.

[1] The SDCP grouped the counties geographically into clusters to assist programme management.

One of the major benefits of the criteria-based targeting exercise was that, based on the insight of the experts, we excluded divisions where other policies or actions may have had differential effects over time. This was especially beneficial since this is an ex-post evaluation for a long-running programme. Specifically, we excluded divisions where county-level policies provided a high level of support, where other dairy programmes provided a high level of support, and where 2007 post-election violence was extreme.

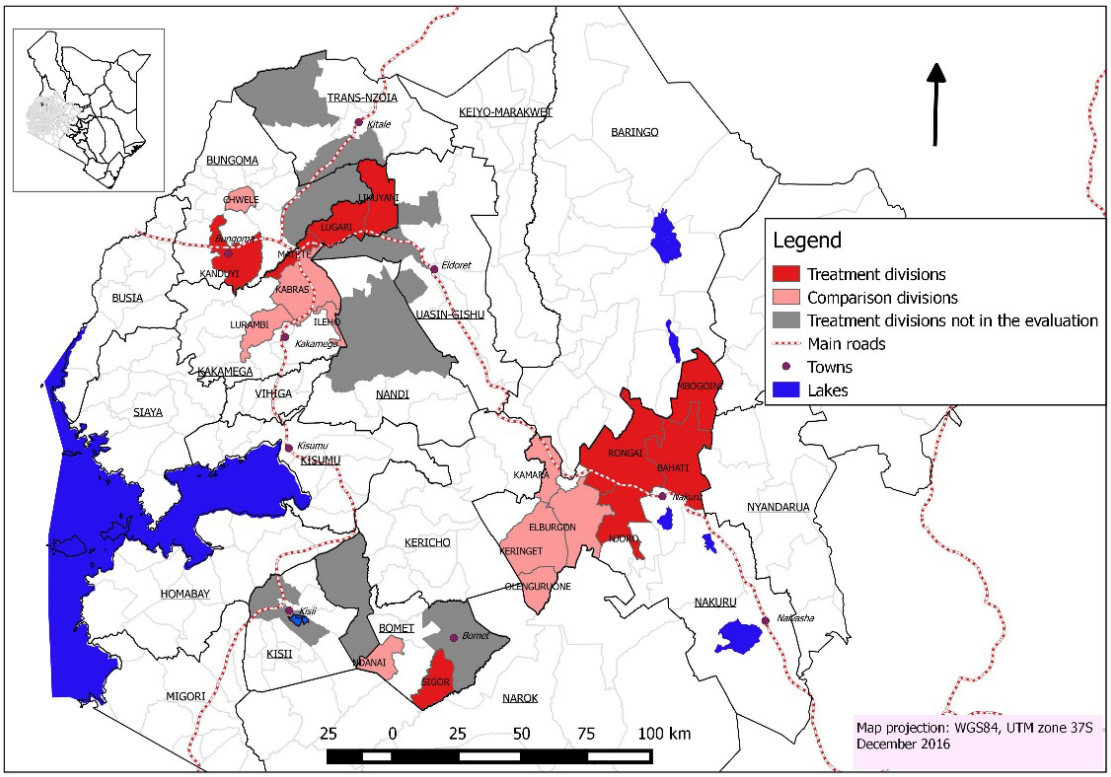
We cross-referenced the sites recommended by the experts from the criteria-based targeting exercise, with the results of the division-level propensity score matching from the observation-based targeting exercise to arrive at the final sample. The results of the two-part targeting exercise are detailed in Table 1 and further visualised in Figure 2.

Table 1: SDCP divisions and evaluation study areas

County	SDCP divisions	SDCP division(s) included in the evaluation	Comparison area divisions included in the evaluation
Cluster: Bomet, Nyamira and Central Kisii			
Bomet	Bomet Central Longisa Sigor	Sigor	Ndanai
Nyamira	Borabu Ekerenyo Nyamira	N/A	N/A
Central Kisii	Keumbu Mosocho Suneka	N/A	N/A
Cluster: Nakuru			
Nakuru	Rongai	Rongai	Kamara
Nakuru	Njoro	Njoro	Keringet
Nakuru	Subukia-Kabazi: Mbogoini and Bahati	Subukia-Kabazi: Mbogoini and Bahati	Elburgon and Olenguruone
Cluster: Lugari, Bungoma, Nandi, Trans Nzoia and Uasin Gishu			
Lugari	Lugari	Lugari	Lurambi[2]
Lugari	Likuyanki	Likuyanki	Ileho
Lugari	Matete	Matete	Kabras
Bungoma	Kanduyi Ndivisi Tongaren	Kanduyi	Chwele
Nandi	Kabiyet Kapsabet Kilibwoni Kosirai	N/A	N/A
Trans Nzoia	Endebess Kiminini	N/A	N/A
Uasin Gishu	Kapsaret Soy Turbo	N/A	N/A

[2] The comparison area for Lugari division was initially Ikolomani. However, when data collection began, we learned that Ikolomani had only a farming cooperative, not farming groups. So, we replaced Ikolomani with Lurambi, the next best match in terms of propensity score.

Figure 2: Evaluation treatment and comparison divisions and non-evaluation treatment divisions



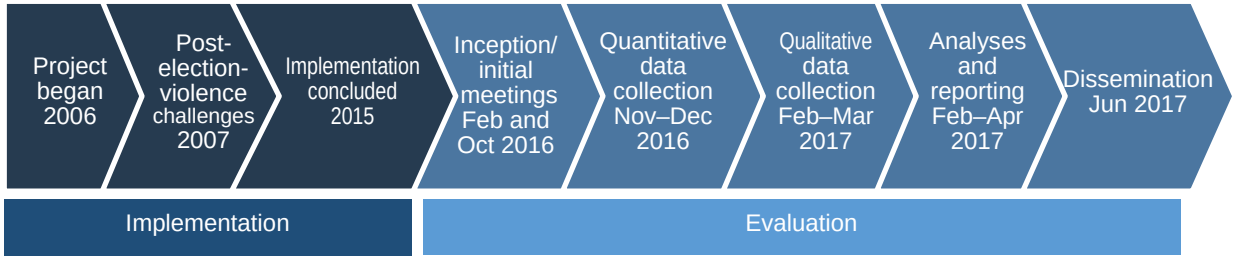
Since the evaluation sampled from all three geographic clusters of programme implementation, the results should be generally representative of the SDCP population. That is, the results should apply to the average dairy farmer who is located in areas with medium to high levels of milk production and poverty. However, because we excluded certain divisions that would have posed difficulties in forming our study group, our sample lacks representativeness on certain dimensions.

Specifically, the study did not include areas where there was a high level of support provided by other dairy-focused programmes, where the aftermath of the 2007 election was severe, where large-scale farmers were more common than smallholders, where farmers were less focused on dairy farming, and where farmers were located far from each other. Nevertheless, the observation-based and criteria-based targeting exercises ensured that we had comparison areas that are similar to the treated areas.

However, there are some limitations to the external validity of the study sites. Specifically, as indicated by the upper left corner of Figure 1, implementation of the SDCP covered districts/counties in the western region of Kenya. Thus, the results do not necessarily extend to other regions of Kenya where dairy farming is a less central focus of smallholder farmers. However, lessons may well be drawn for other East African countries with similar ecological and socio-economic characteristics as those found in western Kenya, including those in the central highlands of Kenya, as well as the highlands of Uganda, Rwanda, Tanzania, and to a lesser extent, Ethiopia (Herrero et al. 2014, Global Environmental Change).

4. Timeline

The following timeline depicts the stages of implementation and evaluation. As an ex-post evaluation, all evaluation activities occurred after implementation of the programme.



5. Evaluation: design, methods and implementation

5.1 Institutional review board

AIR is registered with the Office of Human Research Protection as a research institution (IORG0000260) and conducts research under its own Federalwide Assurance (FWA00003952). AIR's Institutional Review Board (IRB) (IRB00000436) reviewed our procedures for minimising the risks to participants, along with the instruments and protocols. To ensure ethical research, RSA read a consent statement that explained the purposes of the research and the expected duration of the subject's participation, and which described the procedures to be followed.

5.2 Quantitative design

It was necessary to establish a clear counterfactual to conduct a valid assessment of the impact of the SDCP on smallholder farmers. To address the question of what would have happened to programme participants had they not received the intervention, we used a rigorous quasi-experimental methodology. An experimental design was not possible because of the ex-post nature of the evaluation. However, estimating programme impacts by comparing a treatment group with a non-experimental comparison group may be biased because participants self-select into the programme, or because implementing partners specifically target those beneficiaries that are more likely to experience the largest programme impacts.

Within the context section above, we describe how the SDCP targeted programme beneficiaries and how we used that information and observation- and criteria-based targeting to select the study sites. First, we used existing administrative information from the original programme targeting to identify (through propensity score matching at the division level) similar non-programme areas that were not affected by the SDCP due to capacity constraints (i.e. observation-based targeting). Second, we discussed the filtered control communities with local experts and stakeholders to determine which of the preselected control areas were more similar to treatment communities at project inception (i.e. criteria-based targeting).

After matching comparison areas to treatment areas for the study sites, RSA, with oversight from LEAD Analytics, conducted a comprehensive dairy survey to a sample of

2,562 dairy farmers, which covered key information needed to map out the causal chain among inputs, activities, outputs, outcomes and impacts, as well as the underlying assumptions of the SDCP. The sample of 2,562 was split between 1,297 SDCP beneficiaries (from 95 dairy groups) and 1,265 matched comparison farmers (from 89 dairy groups).

To capture key time-invariant characteristics and retrospective information on dairy farming, we included a short filter questionnaire at the beginning of the household survey to improve the matching process between each of households in the treatment group to a similar farm household in a comparison area. Only farm households that were eligible as potential comparisons based on predetermined variables were part of the sample. The idea behind the filter questionnaire was to mimic the selection mechanism that SDCP staff used when defining the dairy groups that later joined the programme.

The identification strategy proposed to estimate the causal effects of the programme relies on the doubly robust estimator developed by Robins and Rotnitzky (1995), Robins et al. (1995), and Van der Laan and Robins (2003). At household level, the proposed approach combines regression and propensity score matching methods in a three-step approach to estimate treatment effects.

In the first step, a treatment model is defined that explains the probability of programme participation. From this step, inverse-probability weights are derived from the estimated propensity score. Second, using the estimated inverse-probability weights, weighted regression models are fit for the outcome equation for each treatment level and obtained the treatment-specific predicted outcomes for each subject. Last, means of the treatment-specific predicted outcomes are computed, and the difference of these averages provides the estimate of the average treatment effect of the programme. Intuitively, weighting can be interpreted as removing the correlation between the treatment condition and other covariates that may be correlated with treatment, and regression as removing the direct effect of such variables on the outcomes of interest (Imbens and Wooldridge 2009).

This approach assumes that programme participation is exogenous to potential outcomes conditional on observable characteristics; that is, that there is no selection bias due to unobserved characteristics, and that the observable characteristics we capture determine programme participation. Due to the unobservable nature of these potential additional characteristics, this assumption is untestable. Nevertheless, we employed a series of strategies to reduce the potential threat of the impact estimates being driven by unobserved characteristics of programme participants.

Specifically, in addition to replicating the division selection process that was conducted to determine programme placement, we used a filter questionnaire to replicate the selection of potentially eligible dairy groups, and to collect numerous covariates as controls that are good predictors of programme participation. Several authors have argued that social programmes can be evaluated using matching methods, as long as there is access to a rich set of variables that determine programme participation, and the non-experimental comparison group is drawn from the same local region as participants (Heckman et al. 1997; Heckman et al. 1998). We were confident that our proposed empirical strategy would allow us to estimate the causal effect of the SDCP on smallholder dairy farmers.

5.3 Sample size determination

We determined that a sample of 2,500 farmers, split evenly between the treatment and comparison groups, would be sufficiently large to detect meaningful programme effects. To calculate the sample size, we conducted a number of power analyses based on existing farm-level data for SDCP beneficiaries that were collected as part of a 2014 survey by Capital Guardian Consulting that the SDCP commissioned. These calculations are described in more detail in Appendix A. The calculations account for the fact that the dairy group serves as the clustering variable.

We proposed to collect information for 154 dairy groups (77 treated and 77 comparison). For each dairy group, we proposed to randomly select 15 households. In practice, we conducted a full survey instrument to the sample of 2,562 observations (1,297 treatment and 1,265 comparison dairy farm households), coming from 183 dairy groups (95 treatment and 89 control). In the survey, we collected key information to map out the causal chain among inputs, activities, outputs, outcomes and impacts, as well as the underlying assumptions.

5.4 Sampling design

To increase comparability beyond the efforts we took for the study site selection, we included a short filter questionnaire at the beginning of the household survey to capture key time-invariant characteristics and retrospective information on dairy farming to improve the matching process between each of households in the treatment group to a similar farm household in a comparison area. We included only dairy farmers who met characteristics that should have increased comparability between the treatment and comparison groups.

The filter questionnaire ensured that only those potential comparison group farmers who were most similar to the treated farmers remained in the final sample. Furthermore, this procedure represented an efficient use of project funds and helped reduce the time burden placed on the farmers. By including a filter questionnaire that eliminated dissimilar farmers, we ensured that project resources spent on the full household survey were allocated to those farmers most relevant for inclusion in the study. Additionally, these procedures reduced the time burden associated with the survey, since we administered the survey to relevant farmers only.

5.5 Data collection

AIR, in conjunction with Lead Analytics, designed the quantitative questionnaire, adapting questionnaires used by ILRI in Tanzania and the World Bank Living Standards Measurement Study – Integrated Surveys on Agriculture in Malawi. The questionnaire was designed to collect detailed data about milk production, cost and sales to generate information on net milk income and milk sales, which are two primary outcomes of interest to assess project impact.

The project also aimed to reduce seasonality of milk production, so that net incomes would be higher and less variable throughout the year. Thus, the questionnaire also collected data on practices (such as second-season fodder grass production) associated with less pronounced seasonality in milk production. Finally, the hoped-for impacts

included increased food security. In part, lower seasonality should contribute to more smooth consumption patterns throughout the year, reducing or eliminating the lean season. Second, while a full-scale consumption module was not included, we did include a module to capture dietary diversity, based on recent recall data.

Building on evidence from the literature, the questionnaire also included sections to recover information on the most important control variables at the household level, in order to improve precision of estimating project impact. These included basic household demographics and wealth variables, landholdings, access to extension and other sources of information, density of social networks, etc.

Importantly, we also designed a dairy group questionnaire. The functioning of dairy groups (i.e. structure, conduct and performance) is likely to have had a strong impact on the ability of households to benefit from project activities, many of which were carried out through the dairy group leadership. Indicators of dairy group performance could serve as controls, and also provide valuable additional insights to feed into future project designs.

The dairy group questionnaire also included a module on the history of presence of other development projects in addition to the SDCP, which could prove to be useful control information, as well as basic information on community characteristics. In control villages, where no dairy groups are currently functioning, the community-level questions were addressed to village leaders.

The paper questionnaire was translated to tablet, primarily by our local survey partner RSA, with inputs from Lead Analytics during the training and piloting phases. The questionnaires were conducted in the field by RSA between 15 November and 20 December 2016. This time frame corresponded to the secondary rainy season in Kenya, which happens for a few weeks in November and December and is followed by a dry season of hot weather until March. The heaviest rainy season in Kenya occurs in late April, May and early June.

The survey time period closely corresponded to the end of harvest period for the main rainy season.[3] A LEAD Analytics field manager oversaw the training and piloting of data collection, and after coordinating with the other AIR and LEAD Analytics team members, provided daily feedback to the enumerators and survey programmers throughout both processes.

5.6 Strategies to avoid biases

Training for our quantitative and qualitative data collection was key to ensuring that interviewers understood the study, interview protocols, interviewing techniques, and the importance of understanding questions exactly as written and recording responses

[3] In Kenya, the following crops are produced during the main rainy season: sorghum, which is harvested by the end of September; beans and millet, which are harvested by the end of October; and maize and wheat, which are harvested by the end of November. Barley, maize, millet, sorghum and beans are also produced during the short rainy season and harvested at the end of March. See Kenyan Crop Calendar available at FAO Global Information and Early Warning System Country Brief on Kenya [URL: http://www.fao.org/giews/countrybrief/country.jsp?code=KEN].

exactly as stated – all of which improve data quality and reduce bias. In addition, the training covered administering and obtaining consent from every participant to ensure compliance with AIR's IRB. The training covered the following topics:

- overview of the SDCP
- overview of AIR's evaluation of SDCP
- ethics, consent and confidentiality
- basics of qualitative research (interview protocols and observations)
- professional conduct
- data security
- in-depth practice of all protocols, including translations and role playing
- team debriefs and protocol revisions
- field plans and logistics.

We further reduced biases by choosing comparison areas, as opposed to comparison households. The use of comparison areas helped overcome concerns about not capturing the true programme impacts due to spillover effects. If farmers in the comparison areas were outside the SDCP's catchment area, the benefits would be less likely to flow to the comparison areas. We investigated the validity of this requirement through the farm-level surveys by asking farmers in comparison areas how much they knew about IFAD's SDCP and other agricultural development programmes.

Given that numerous development and other agricultural programmes have been implemented in Kenya, it was important that the comparison and treatment groups had the same level of exposure to these programmes over time, with the exception of the SDCP. Many of these other development and agricultural programmes have been implemented at the county level. By choosing comparison areas that were within the same counties as the treated DCAs, we could account for exposure to other dairy programmes. Although we chose comparison areas within the same counties, we made sure not to choose comparison divisions that directly neighboured a treatment division. Ensuring that treated and comparison areas were sufficiently spaced apart minimised the risk of spillovers.

An obvious concern when evaluating the programme was that due to a sense of reciprocity (i.e. Hawthorne Effect), respondents in the treatment group may have offered answers that they believed the research team was seeking. We structured the questionnaire in such a way that the questions related to the programme intervention were non-conspicuous and occurred after questions related to measurement of important outcomes, so as to delay activation of potential Hawthorne effects. In addition, we do not believe this evaluation exhibited John Henry effects, wherein non-treated farmers react to overcome the disadvantage of being in the control group, because we did not expect that comparison farmers knew much about the SDCP, since they were drawn from different geographical areas.

5.7 Qualitative design

Addressing the research questions required a combination of qualitative and quantitative methods. In qualitative research, questions – and the responses they elicit – tend to be discursive and descriptive, while their analysis privileges explanation and interpretation

over quantification. In general, qualitative approaches allow researchers to explore and understand the experiences, opinions and perspectives of their informants in greater depth than that offered by quantitative approaches.

In turn, the use of qualitative approaches entails sacrifices in terms of generalisability and comparability – areas in which quantitative methods excel because of their use of large and probabilistic samples. Samples chosen for qualitative studies are often non-randomised (or purposively selected) and always smaller: ‘There is growing evidence that 10–20 knowledgeable people are enough to uncover and understand the core categories in any well-defined cultural domain or study of lived experience’ (Bernard 2011, p.154).

We drew on qualitative methods to augment the quantitative surveys by capturing interaction among complex and changing contextual factors that could influence the impact of the SDCP, and evaluated programme fidelity. The process component also aimed to assess gaps in implementation that may have affected impacts, and how dairy farmers may have changed their practices based on what they learned with regard to Components 1–3. The contextual information we obtained through the qualitative information helped to clarify how the programme impacted individuals, thus contributing to the transferability of study findings to other settings (i.e. external validity).

5.8 Design

We conducted semi-structured interviews with key informants directly involved in the programme at the national, county and sub-county levels. These discussions aimed to clarify uptake of the programme and the coordination of county departments of agriculture with the farmers. Second, we conducted key informant interviews with private service providers who are unrelated to the programme to ascertain how the dairy sector functions outside of the programme and the public sector.

5.9 Qualitative sample

We conducted 12 focus groups using semi-structured protocols with 6–8 farmers each from 12 dairy groups across two counties. We included two counties with three SDCP administrative areas each in the qualitative sample, just in case there were any notable regional or within-county differences by administrative area. We included dairy groups with varying proportions of male and female farmers. The focus group sample was half male and half female, and we separated the groups by male and female farmers to capture a better understanding of experiences by sex. Finally, we included groups who were considered to be high performing and low performing, based on consultations with stakeholders familiar with the dairy groups.

We structured the questions to understand whether all of the groups engaged similarly with the various components, or whether and how some of the components were stronger for some groups than others. To compare these data with how groups functioned in the absence of the programme, we also conducted four FGDs (two male and two female) with farmers in comparison areas. We aimed to cover all components with some piece of the qualitative work (Table 2).

Table 2: Qualitative sampling

Stakeholder group	Geographical level	Number of interviews	Component focus
SDCP Coordinating Unit	National	4	5 (1, 2)
Dairy Training Institute	National	1	3
Kenya Animal Genetic Resources Centre	National	1	3
Department of Livestock	Regional	3	3, 4
Community artificial insemination station	County	Nakuru, 2; Lugari, 1	3
Input supplier	County	Nakuru, 1; Lugari, 1	3, 4
County Programme Coordination Team	County	Nakuru, 2; Lugari, 3	5 (1, 2)
Treatment dairy group chairs	County	5	1, 2
Treatment smallholder dairy groups	County	12 (1 man and 1 woman for 6 dairy groups)	1, 2
Comparison dairy group chairs	County	1	1, 3
Comparison dairy groups	County	4 (1 man and 1 woman for 2 dairy groups)	1, 3

5.10 Methods

To ensure reliability of the findings, the research team employed several analytic methods to systematise the data review and coding: (a) content coding; (b) comparison of findings among researchers; and (c) grounded theory to guide analysis using NVivo qualitative data analysis software.[4]

The first step in analysing qualitative data is to develop a coding structure that helps to systematically categorise information. Researchers 'open code' data to identify primary categories of interest. These categories form the basis for the coding structure that the team uses to categorise raw data from interviews and FGDs into the primary findings.

While categorising the data, researchers applied grounded theory (Glaser and Strauss 2009) to deduct new themes from the findings, rather than testing an existing theory. Combining the use of grounded theory with a rigorous impact evaluation design enabled us to triangulate the research findings. The quantitative research served to test predefined hypotheses, and the qualitative research enabled us to ground new ideas on why the programme positively influenced some, but not all, outcomes of interest among respondents' real-life experiences.

After coding, the research team attempted to quantify the data where applicable. This method helped characterise the prevalence of responses to deduce which themes were common and which were outliers. It is important to note, however, that qualitative data are not best analysed using a systematic count of opinions.

[4] QSR International Pty Ltd. Version 10, 2012.

5.11 Implementation: data collection and strategies to avoid biases

In general, qualitative approaches allow researchers to explore and understand the experiences, opinions, and perspectives of their informants in greater depth than quantitative approaches allow. The use of qualitative approaches entails sacrifices in terms of generalisability and comparability – areas in which quantitative methods excel because of their use of large and probabilistic samples. Interviewers were clear on these limitations, as well as their neutral role, which aimed to limit courtesy and social desirability bias (a situation in which a respondent gives an answer that he or she feels the interviewer wants to hear).

Data collection occurred over a period of two weeks between 27 February and 10 March 2017. The evaluation team trained data collectors before they went into the field. We held a plenary training to introduce participants to the study, the supervision process, logistics, interviewing techniques, and the interview and observation protocols, and also to conduct mock interviews. During training, the research team improved the items on the interview protocols by incorporating input from local interviewers. During the initial days in the field, interviewers noted challenges in administering the surveys and interview protocols, which we discussed during debriefings during the first phase of fieldwork.

The team collected data using notes and digital recordings. All of the interviews and FGDs were transcribed. The team ensured complete anonymisation and protection of confidentiality for research participants.

5.12 Limitations of the study

As noted before, the methodology employed with the quantitative design assumed that there was no selection bias due to unobserved characteristics, and that the observable characteristics we captured determined programme participation. However, due to the unobservable nature of these potential additional characteristics, this assumption was untestable. Additionally, while our qualitative research provided insights into the uptake rates of the programme, we could not determine these rates quantitatively, because tracking down targeted farmers from 2006 was not feasible. This limitation prevented us from analysing what impact, if any, the programme had on the likelihood of a dairy farmer in 2006 remaining one to the present day.

6. Programme or policy: design, methods and implementation

The SDCP is a joint programme between the GoK and IFAD that commenced in July 2006 and will be completed by September 2019. The SDCP was designed to reach dairy groups engaged in milk production in nine milk-producing counties. The overall goal of the SDCP was to increase the income of poor rural households that depend substantially on the production and trade of dairy products for their livelihoods.

The SDCP was implemented through various interrelated components. First, the SDCP provided training on organisational, managerial and enterprise skills (e.g. bookkeeping, accounting and financial planning) to farmers. Second, the programme targeted household production by aiming to enhance dairy farming productivity and reduce production costs through grants, trainings, field days and demonstrations. Last, the

SDCP aimed to strengthen market linkages for small-scale milk producers by improving road infrastructure and conducting additional trainings on milk-handling practices and value addition opportunities. The project operated through dairy groups and was intended to reach 24,000 smallholder dairy-farming households, with members participating in 600 dairy groups across 9 milk-producing counties.

Given that the implementation began in 2006, we did not have a monitoring system to track implementation rollout. Similarly, since the evaluation was an ex-post evaluation, we did not provide incentives for participation in the study groups. For the quantitative data collection and key informant interviews, we did not provide any compensation or incentives for completion of the survey or interviews. For participation in the qualitative FGDs, we provided a transport allowance and snack.

7. Impact analysis and results of key evaluation questions

7.1 Descriptive statistics and quality of counterfactual

In this section, we present some of the key explanatory variables collected for the study. These variables help us to understand the context in which the SDCP was implemented and serve as controls in the econometric models. This section also describes the outcomes of interest and presents the impact results.

ILRI collected the variables used for the matching procedure at the division level for programme placement before 2005. These included milk density, proportion of poor households, and travel time to the nearest urban centre, which we have included in Table 3. The original ILRI data also included the number of households and area, although the ILRI final targeting exercise only used milk density and proportion of poor households.

As shown in Table 3, milk density in the study area is on average 83,000 litres/km^2/year. The difference in milk density between the SDCP areas and control areas is not statistically significant. From examining the proportion of poor households in the division, we see that more than half of the households in our dataset are poor. The proportion of poor people in the SDCP households and control households is similar, at 53 per cent. In terms of distance, SDCP households travel 0.39 hours to the nearest urban centre, as do households in the comparison group.

The averages for these three key results are similar for the treatment and control groups. More specifically, Table 3 also presents the results of a linear model for the probability of being in the treatment group as a function of the three division-level characteristics used for programme placement. This regression is conducted at the dairy group level and includes county fixed effects, which is the administrative level at which the initial matching of divisions was conducted.

The results show that none of the three placement variables is a good predictor of the probability of being in the treatment group. In fact, we are not able to reject the hypothesis that the three variables are jointly equal to zero using an F test (p-value = 0.86). These results provide evidence that we are able to find divisions in the same counties where the SDCP operates that could have been chosen for the programme

based on the three placement characteristics, but were not chosen due to programme capacity constraints.

Table 3: Pre-programme, division-level characteristics used for matching (N=183)

Placement variables	Mean control	Mean treatment	Coefficient	Standard error	p-value
Milk density (litres/km^2/year)	85,389	81,099	-0.000008	0.000009	0.398
Proportion of poor people	0.53	0.53	0.15	3.49	0.967
Travel time to nearest urban centre	0.39	0.40	-0.55	0.92	0.557

Note, however, that finding control divisions that were similar to SDCP divisions did not guarantee that the individual dairy groups within these divisions were similar. To increase comparability beyond the study site selection, we included a short filter questionnaire at the beginning of the household survey to capture key time-invariant characteristics and retrospective information on dairy farming to improve the matching process between each of households in the treatment group and a similar farm household in a comparison area.

In Table 4 we specify a linear probability model to assess to what extent the treatment and control households differ in terms of key household and agricultural characteristics.

In terms of household demographics, the data show that the typical household spoke local languages (only 8 per cent of the households spoke English), and had around three people in the working age group (people aged 14–65 years old) and 0.5 children aged five years of age and younger. Also, on average, the SDCP household head was 52 years old and had 13 years of education. Lastly, about 24 per cent of households in our data had a female head of household. The percentage of female-headed households was 11 percentage points higher in the SDCP group than in the control group, a difference that was statistically significant.

Regarding the household socio-economic and agricultural characteristics, the results show that there were no statistically significant differences between treatment and control households in terms of land size, slope of farm plots, access to irrigation and area cultivated. The average household in the sample had 0.19 acres, with a legal title; 87 percent of the farms were either flat or had a slight slope, and did not have access to irrigation.

In turn, there are some variables where treatment and control households exhibited differences that were statistically significant. A one standard-deviation increase in the indices of consumer durables and agricultural implements increased the probability of being in the treatment group by approximately 2 percentage points. Also, treatment households were 10 percentage points more likely to have cultivated a crop in the primary agricultural season.

There were some differences in the breed composition of the division where the household was located as measured in 2005.[5] An increase of 100 head of zebu in 2005 on the division where the household was located decreased the probability of being in the treatment group by 10 percentage points. Alternatively, an increase of 100 head of crossbred cattle increased the probability of being in the treatment group by 10 percentage points. There were no differences in 2005 between treatment and control households in terms of the presence of grade cattle in their corresponding divisions.

Lastly, treated and comparison households have similar probabilities of having received other agricultural or livestock development programmes.

Table 4: Determinants of treatment probability (N=2558)

Independent variable	Mean	Coefficient	SE	p-value
Household demographics				
Main language is English = 1	0.08	-0.011	0.028	0.70
No. people in working-age group	2.94	-0.005	0.006	0.43
No. children five years of age and younger	0.48	0.017	0.012	0.14
Age of household head	52.3	0.003	0.001	0.00
Max. years of schooling	13.0	0.003	0.004	0.46
Female head of household = 1	0.24	0.112	0.025	0.00
Income and agricultural characteristics				
Consumer durables index (SD)	0.00	0.018	0.010	0.08
Ag. Implements index (SD)	0.00	0.019	0.008	0.01
Total size of landholdings (acres purchased, with title)	0.19	0.015	0.027	0.59
Farm is flat or with slight slope = 1	0.87	-0.008	0.028	0.77
Plot with any system of irrigation = 1	0.02	-0.082	0.075	0.28
Plot with any irrigation structure = 1	0.02	-0.057	0.079	0.47
Household cultivated any crop in primary season	0.88	0.100	0.033	0.00
No. zebu cattle in division in 2005	1071.1	-0.001	0.000	0.00
No. cross cattle in division in 2005	1388.2	0.001	0.000	0.00
No. grade cattle in division in 2005	959.1	0.00003	0.000	0.83
Development projects in community				
Any agriculture-based project other than SDCP = 1	0.55	-0.002	0.088	0.99
Any livestock-based project = 1	0.12	0.078	0.090	0.39

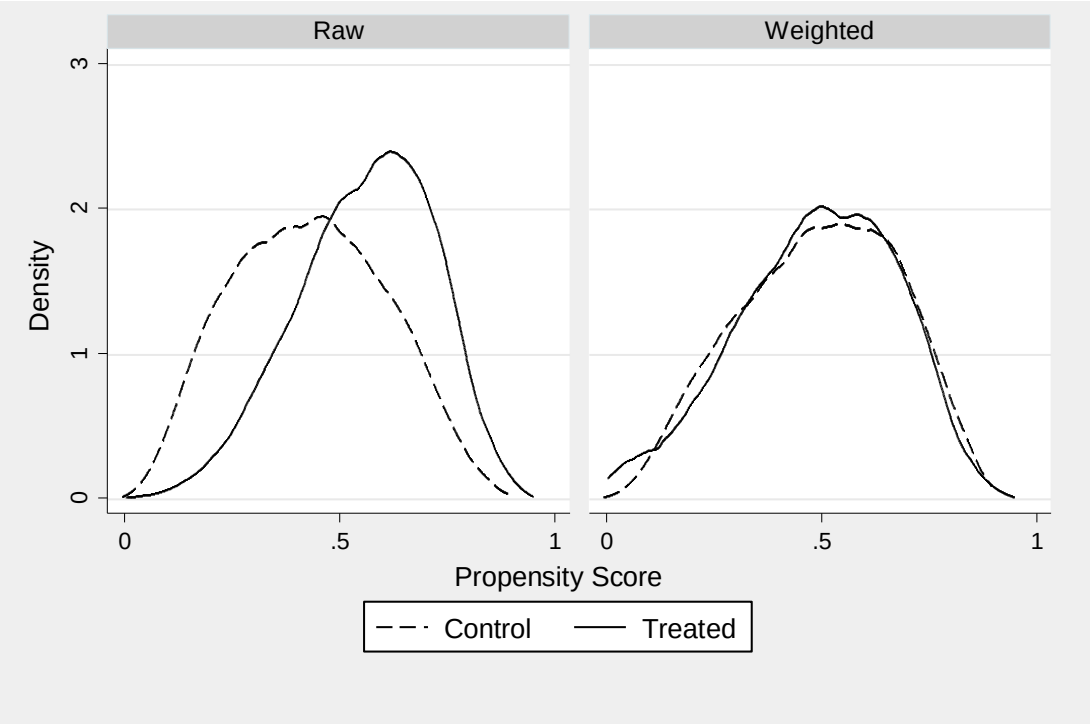
Notes: This table shows the results of a linear model for the probability of being in the treatment group as a function of household, agricultural and community characteristics. The table shows: the mean for each explanatory variable; change in the propensity of being in the treatment group after a change of one unit on a given characteristic; standard error of each estimate; and corresponding p-value.

Overall, the results show that, with the exception of a few variables, the treatment and control households had very similar household and socio-economic characteristics. Nevertheless, some of the reported differences in observable characteristics could be an

[5] ILRI also constructed these variables in 2005 at the division level as part of the programme placement exercise, but they were not used as selection variables.

indication that there were differences in unobservable characteristics between the two groups that may ultimately have biased programme impacts. Indeed, the balance plot in Figure 3 below shows there was a skew towards relatively high propensity scores for the treated households, which motivated the use of the Inverse-Probability-Weighted Regression Adjustment (IPWRA). It is reassuring for the identification strategy that the weights produced very similar densities of the propensity score.

Figure 3: Propensity score kernel densities before and after weighting



Moreover, we also conducted balance tests for each covariate considered in the treatment equation, where each observation was weighted by an inverse function of the estimated propensity score. As shown in Table 5, all the covariates considered were balanced between the treatment and comparison groups when the full sample was considered. We also conducted covariate balance tests by strata of the propensity score to check whether the different covariates were balanced at different points over the distribution of the propensity score.

For this analysis, we considered seven blocks, or strata, of the propensity score to ensure that the mean propensity score was not different for treated and control households in each block. Given that our treatment equation included 20 covariates, we conducted a total of 140 tests (=7*20) of the balancing property. Of those, only 5 tests (3 per cent) rejected the hypothesis of treated and control households having the same mean for a given variable within a given block. Overall, the results indicated there was a good covariate balance after implementing a strategy based on inverse-probability weighting.

Table 5: Covariate balance tests over treatment groups with inverse-probability weighting

Dependent variable	Mean	Coefficient	SE	p-value
Household demographics				
Main language is English = 1	0.08	-0.015	0.017	0.365
Number of people in the working age group	2.94	-0.105	0.098	0.285
Number of children five years of age and younger	0.48	-0.004	0.032	0.903
Age of household head	52.3	0.002	0.012	0.838
Max. years of schooling of household member	13.0	-0.140	0.163	0.390
Female head of household	0.24	0.002	0.020	0.914
Income and agricultural characteristics				
Consumer durables index (SD)	0.00	-0.083	0.088	0.344
Ag. Implements index (SD)	0.00	-0.069	0.086	0.422
Size of landholdings (acres purchased with title)	0.19	0.004	0.016	0.782
Household cultivated any crop in primary season	0.88	-0.005	0.015	0.739
Household cultivated any crop in secondary season	0.02	-0.007	0.022	0.737
Area cultivated in primary season (log of ha)	0.22	-0.013	0.021	0.552
Area cultivated in secondary season (log of ha)	0.05	-0.011	0.021	0.599
Farm is flat or with slight slope = 1	0.87	-0.006	0.015	0.683
Any system of irrigation = 1	0.02	-0.009	0.012	0.430
Any irrigation structure = 1	0.02	-0.009	0.011	0.444
Development projects in community				
Community has any agriculture-based project other than SDCP =1	0.55	-0.007	0.022	0.768
Any project in community is focused on livestock	0.12	-0.014	0.017	0.412

Notes: Each row in this table corresponds to an OLS regression of the specified variable on the treatment indicator. The table shows: the mean for each dependent variable; difference in means between the treatment and control groups (coefficient); standard error for each difference; and corresponding p-value.

Table 6 indicates that more than half of the households in our dataset reported that there were other agriculture-based projects in their communities other than SDCP. Most of those projects were agricultural ones; those that focused on livestock were only a small portion of the projects available.

Table 6: Availability of other agricultural and livestock support programmes

Characteristics	SDCP	Control	All
Community has any agriculture-based project other than SDCP	0.56	0.54	0.55
Any project in community is focused on agriculture	0.42	0.45	0.43
Any project in community is focused on livestock	0.15	0.09	0.12

7.2 Mixed methods analysis (quantitative and qualitative)

7.2.1 Methods: regression techniques

As indicated above, at the household level the identification strategy that was used to estimate SDCP impacts combined regression and propensity score methods in a three-step approach to estimate treatment effects. In the first step, we estimated the probability of programme participation through a logit or probit such as:

$$T_i = \Phi(\boldsymbol{X}_i \cdot \boldsymbol{\beta}' + \varepsilon_i)$$

where T_i is a dummy for having received the programme, X_i is a vector of individual and division-level characteristics, and ε_i is an error term. The observable characteristics considered in the treatment equation included the 2006 division-level characteristics used for the initial programme placement, namely, milk density (litre/km^2/year), the proportion of poor people in the division, and the travel time (in hours) to the nearest urban centre.

In addition to these variables, we included other 2006 division-level variables constructed by ILRI that accounted for the estimated number of heads from the different breed categories (e.g. local, crossed and exotic), which were a good indication of the division breed composition and good predictors of the ability to increase the prevalence of higher milk-producing breeds.

We controlled for household-level characteristics such as the sex and age of the household head, the head's maximum number of years of schooling, the language spoken at home, the number of working age members, and the number of children under the age of five. We also controlled for indices of consumer durables and agricultural implements as a way to control for household income differences, as well as dummies for the existence of other agricultural and livestock projects (different from the SDCP) in recent years in the village where the respondent lived.

To account for differences in agricultural production in the estimation of the propensity score, we controlled for a set of agricultural production characteristics such as indicator variables for producing any crops in the primary and secondary rainy seasons; the natural log of the hectares cultivated in each one of the two seasons; dummies for the slope of the main plot being flat or slightly flat; indicator variables for the availability of irrigation in at least one of the plots; and the percentage of plots served by an irrigation scheme. We also controlled for county-level fixed effects to ensure that the comparison of treatment and comparison households was done within a given county.

Then we generated propensity scores for P_i, the probability of receiving treatment, as

$$P_i = \Phi(\boldsymbol{X}_i \cdot \widehat{\boldsymbol{\beta}}')$$

Second, using the estimated inverse-probability weights, weighted regression models were fit for the outcome equation for each treatment level and obtained the treatment-specific predicted outcomes for each subject. Last, means of the treatment-specific predicted outcomes were computed and the difference of these averages provided the estimate of the average treatment effect of the programme. More specifically, we estimated programme impacts by using the *teffects ipwra* in STATA using the *osample*

option. Note that all observations in the original sample were used in analysis, as no observations violated the overlap assumption.

7.3 Intermediate outcomes

7.3.1 Improved animal management

We start the presentation of the programme impacts by looking at the results on animal management. In Table 7 we show the results for the impact of the programme on feeding and water practices farmers used with their cattle. Although the programme did not have a statistically significant impact across all dimensions of feeding and water practices, positive impacts existed related to the grazing practices, keeping practices and feeding of concentrate feeds.

Specifically, we found that SDCP increased the probability that farmers practised zero grazing with their cattle by 8 percentage points. SDCP staff stated that practising zero grazing was one of the key messages that the programme delivered to farmers to control how much cows ate and the nutritional value of what they ate, and to reduce losses in potential milk production from searching for food over large areas.

Improved animal management was the most frequently discussed benefit of the SDCP among FGD participants. One farmer described the changes she made in animal management: ‘Initially I would send the children to go and graze, [but] I now practise zero grazing. I realised I used to lose a lot on milk yields. At least I can produce more milk, I treat [the cows] when they fall sick, clean them and ensure they are vaccinated.’

We also found that the SDCP increased the probability that farmers kept their cattle in a paddock that had a stall floor made of concrete by 5 percentage points. This increase was promising since concrete floors are easier to clean, which reduces the chances that the milk could be contaminated, which may have had a positive effect on the selling price charged to some customers.

We also found that the SDCP increased the probability that farmers fed cattle concentrate feeds and mineral supplements by 13 percentage points. The programme also showed a positive impact on the probability of feeding cattle off the ground in a trough when milking, which helps protect food from contamination and promotes a more natural grazing posture. These results suggested that farmers were receptive to programme education regarding nutritional content and feeding practices, which may have had a positive effect on milk production.

Table 7: Impacts on feeding practices (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Practise zero grazing	***0.08	0.09	0.17
	(0.02)		
Paddock/boma/stall floor made of concrete	***0.05	0.07	0.12
	(0.02)		
Purchase protein-rich fodder	0.01	0.02	0.03
	(0.01)		
Feed cattle crop residues	-0.00	0.84	0.83
	(0.03)		
Feed cattle concentrate feeds/mineral supplements	***0.13	0.66	0.79
	(0.04)		
Feeding off the ground in trough when milking	***0.12	0.18	0.29
	(0.03)		
Vary feeding by lactation stage	0.03	0.55	0.58
	(0.04)		
Experience shortage of feeds from their farm	-0.01	0.74	0.73
	(-0.03)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

The SDCP also led to an improvement in breeding services. Specifically, the SDCP increased the propensity of households monitoring their cattle on a regular basis and the use of AI services.

Table 8 shows that treatment farmers were 7 percentage points more likely to report that they monitored their cattle regularly. Similarly, treatment farmers were 12 percentage points more likely to report that they had used the AI service in the past 12 months. However, we did not find any statistically significant differences between the treatment and comparison groups related to availability and use of own bull service for breeding.

Table 8: Impacts on access and use of livestock technology (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Household monitors cattle regularly	***0.07	0.88	0.94
	(0.02)		
Technology for own bull service is available	-0.02	0.29	0.27
	(0.04)		
Own bull service used in the last 12 months	-0.01	0.13	0.12
	(0.03)		
Technology for AI service is available	0.05	0.60	0.65
	(0.04)		
AI service used in the last 12 months	***0.12	0.24	0.36
	(0.03)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

The positive quantitative findings related to access to and use of AI were especially promising, considering that many farmers discussed ongoing challenges with AI. The use of AI seems to have been a high cost to farmers, and misconceptions about the purpose and potential of AI services meant that farmers could not make informed decisions in these transactions. Many farmers seemed to be unclear on the proper timing for AI, and a few farmers thought they could pay more for the AI provider to select the sex when they inseminated. Other challenges were that farmers did not believe that government-subsidised AI services provided the best breeds; and that they had to purchase AI services too often because they were often unsuccessful.

The programme had positive impacts related to animal health services, which were concentrated on improved access to and use of vaccination services and curative treatment services. Table 9 shows that the SDCP increased the probability that treatment farmers had access to and used vaccination services by 25 percentage points. Similarly, SDCP farmers were 10 percentage points more likely to have access to curative treatment services and 7 percentage points more likely to use them than farmers in the comparison group. However, the programme did not seem to have changed access to other health services such as deworming (anthelmintics) or tick control services.

Table 9: Impacts on animal health services (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Had access to anthelmintics	0.02	0.90	0.92
	(0.02)		
Used anthelmintics	0.02	0.89	0.91
	(0.02)		
Had access to tick control service	0.00	0.89	0.89
	(0.03)		
Used tick control service	0.00	0.88	0.89
	(0.03)		
Had access to vaccination service	***0.25	0.46	0.71
	(0.04		
Used vaccination service	***0.26	0.43	0.68
	(0.04)		
Had access to curative treatment service	***0.10	0.21	0.30
	(0.04)		
Used curative treatment service	**0.07	0.21	0.28
	(0.03)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

7.3.2 Improved access to extension and information

The SDCP was implemented through several components, whereby the dissemination of information on different areas of the productive process were largely conducted through extension services and trainings of dairy group members. In particular, the SDCP provided training on organisational, managerial and enterprise skills (e.g. bookkeeping, accounting and financial planning) to farmers. In addition, the programme also conducted trainings, field days and demonstrations specifically designed to enhance dairy farming productivity and reduce production costs. In this section, we assess the effect of the programme on making these services available, and to what extent the information provided was adopted.

Table 10 presents the results on the reported availability of extension visits, field days and demonstrations. All these outcomes consistently show that farmers in SDCP areas had a greater access to these types of services. Programme farmers were, on average, 12–18 percentage points more likely to report that the different types of extension services were available in their villages.

Table 10: Impacts on availability of cattle extension services (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Extension visits available	***0.14	0.19	0.33
	(0.03)		
Field days available	***0.18	0.18	0.36
	(0.04)		
Demonstrations available	***0.12	0.14	0.27
	(0.03)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

We also investigated if the programme had any effects on the probability of receiving information on specific aspects of the production process, and whether farmers who attended those trainings adopted the recommended practices. The largest programme effects were seen on receiving information on general livestock practices (15 percentage points); milk processing and quality control (10 percentage points); and fodder establishment and fresh milk marketing (7 percentage points).

However, positive impacts were observed for almost all types of cattle-related topics, which may be a consequence of the programme tailoring specific trainings according to the needs and preferences of each dairy group. The results on adopting specific practices closely resembled the impacts reported in Table 11. Interestingly, SDCP farmers did not report having greater access to information on crop best practices relative to the comparison group, which is reassuring as the SDCP programme did not provide any training on crop practices.

Table 11: Impacts on probability of receiving information from trainings (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Crop best practices	0.04	0.17	0.21
	(0.03)		
Livestock best practices	***0.15	0.22	0.37
	(0.03)		
Improved fermentation of milk best practices	***0.06	0.01	0.07
	(0.01)		
Other milk processing and quality control best practices	***0.10	0.03	0.13
	(0.01)		
Managerial, bookkeeping, accounting and finance	***0.04	0.03	0.07
	(0.01)		
Fodder establishment	***0.07	0.05	0.13
	(0.02)		
Hay making	***0.06	0.06	0.12
	(0.02)		
Silage making	**0.04	0.05	0.09
	(0.02)		
Use of chaff cutter	***0.05	0.01	0.05
	(0.01)		
Conservation (crop residues)	***0.03	0.02	0.04
	(0.01)		
Animal registration	***0.03	0.01	0.04
	(0.01)		
Fresh milk marketing	***0.07	0.02	0.09
	(0.01)		
Value addition marketing (e.g. mala and yoghurt)	***0.05	0.01	0.06
	(0.01)		
Group/cooperative milk marketing	**0.02	0.02	0.04
	(0.01)		
Market information searching	***0.02	0.01	0.02
	(0.01)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

7.3.3 Improved dairy group performance

Following the 'structure-conduct-performance' literature, here we first present results on structure, then conduct, and finally, performance. We surveyed 94 dairy groups in 73 SDCP communities, and 90 dairy groups in 83 control communities; we note here that enumerators surveyed the 'most important' dairy groups within communities. The first difference to note is that there were more dairy group questionnaires implemented in the SDCP than in the control communities. Also, the results we present here are naïve differences between SDCP and control dairy groups. Finally, we only summarise the main results of a larger report on dairy groups; full results can be found in that report.

In terms of basic structure, nearly all dairy groups had an elected chair, and nearly all of the groups (97 per cent) were legally registered and had by-laws or a written constitution. Additionally, there was similar representation of women as respondents to the dairy group questionnaire and in terms of positions held within the dairy group – just over 50 per cent for both the SDCP and control dairy groups.

On the other hand, SDCP groups were more likely to also have an elected secretary and treasurer. We expected that groups with more formalised structures should have performed better in this context, primarily because one of the main goals of the programme was to link the dairy groups to the milk value chain rather than focusing exclusively on intercommunity activities, and also because dues were collected and managed, meaning accountability was important.

Committees may also have improved the performance of dairy groups, enabling subgroup members to focus on specific tasks and enabling members with specific expertise to exploit their comparative advantages within a committee. However, the data indicated that SDCP groups were not more likely to have specific committees overall and were significantly less likely to have financial committees.

Additionally, SDCP groups were more likely to have monthly meetings, but control groups were more likely to hold weekly meetings. More meetings may mean greater opportunities to share information and engage in decision making, but too many meetings may also be inefficient. Thus, evidence on structure is somewhat mixed: there were more elected positions, but fewer committees, and meetings were held less often.

We have only two aspects of conduct: how decisions were made and sources of financing accessed. In terms of decision making, there were very few differences between the SDCP and control dairy groups across decisions on three broad areas: financing, marketing and information acquisition, and dissemination. We allowed for decisions to be made by consensus, by the chair with other key members, and by committee.

The only statistically significant difference was that information decisions for the SDCP group were more likely to have been made by the chair plus other key members. For the control dairy groups, information decisions were more evenly spread across the three decision-making mechanisms. On the other hand, it is interesting to note that while finance and marketing decisions were most often made by consensus in more than 50 per cent of both the SDCP and control communities, consensus-based decision-making was less important for information decisions: between 32 per cent and 36 per cent for the SDCP and control communities, respectively.

Next, we asked about four specific types of financing and a fifth ‘other’ category. Here, the SDCP and control groups exhibited a number of statistically significant differences. The SDCP groups were far more likely to obtain financing from micro-finance institutions (14 per cent versus 1 per cent); from ‘other’ sources that were mainly different types of local credit and savings groups (55 per cent versus 43 per cent); and from commercial banks (5 per cent versus 1 per cent). Both the SDCP and control groups, however, also relied heavily on members’ dues, with 93 per cent stating they relied on member dues for both groups. Overall, both the SDCP and control groups exhibited similar patterns in

decision-making procedures, but the SDCP groups were more likely to obtain financing from a wider range of financial institutions.

Finally, we had a number of measures of dairy group performance including the different types and frequencies of conflict, different types of trainings and various other dairy group activities. With respect to conflicts, we asked about unpaid dues, amount of dues, participation in trainings, financial issues, management issues and other conflicts.

For the most part, the SDCP and control groups faced similar percentages of conflicts, with the exception of unpaid dues. In the SDCP groups, only 21 per cent said there were conflicts over unpaid dues, versus 37 per cent for control groups. This is interesting because the SDCP groups were much less likely to have financial committees, but they were also more likely to have an elected treasurer.

Evidence suggests having an elected treasurer may be more important than having specific financing committees, but we did not ask specifically about how members' dues were collected, which could have shed more light on this issue. Also of interest is that 39 per cent of both the SDCP and control groups indicated that there had been conflicts. Outside of unpaid dues, the most important categories of conflicts included management issues, amount of dues and 'other' issues.

Where we saw the most significant difference between the SDCP and control dairy groups was in trainings made available to dairy group members. Of nine dairy group topics, the SDCP groups stated that at least one training had occurred, which was statistically greater than control groups on all nine topics. The differences were particularly pronounced for the following topics: dairy group management, proposal writing and fodder management.

Finally, we asked about a number of other services the dairy groups provided, including organising educational exchange tours, collecting/sharing milk price data, facilitating links between members and input suppliers, facilitating links between members and milk purchasers, contracting with input suppliers on behalf of members, contracting with milk purchasers on behalf of members, and 'other'.

There was very little difference between the SDCP and control dairy groups on the percentage of groups that offered the different services, except that the SDCP groups were more likely to contract with milk purchasers on behalf of members (11% versus 2%). It is also interesting to note that over 70 per cent of groups (both the SDCP and control groups) arranged educational exchange tours, while less than 30 per cent of groups undertook any of the other services.

Overall, the evidence from the dairy group survey suggested that the SDCP dairy groups functioned and performed a bit differently than the comparison groups. In particular, they were more likely to have elected officials below the chair; more likely to access multiple sources of financing; less likely to face conflicts over collecting dues from members; and more likely to contract with milk purchasers on behalf of members, albeit at a relatively low rate (11 per cent).

The most striking difference, however, was in the number of trainings, provided primarily by the SDCP to dairy group members. This suggests that there is room to improve the performance of the SDCP dairy groups, particularly in terms of the services that the SDCP promotes to reduce smallholders' transaction costs in accessing dairy inputs and milk markets, thereby increasing net revenues accruing to smallholders.

7.4 Main outcomes

One of the key goals of the SDCP was to increase the income of poor rural households that depended substantially on production and trade of dairy products for their livelihoods. To attain this goal, the programme focused on improving productivity, making some inputs more accessible and emphasising the importance of value adding and more reliable trade relations. In this section, we investigate programme impacts on some key final outcomes of the programme.

First, we look at the effects that the programme had on cattle size and composition. As shown in Table 12, we found that, relative to the comparison group, the SDCP farmers not only owned 0.5 more head of cattle (including cows, males, heifers, calves and pre-weaning males), but also had a higher number of cows (that calved at least once), as well as animals they were currently milking. Interestingly, these higher impacts on the number of animals was mostly explained by a higher number of crossbreed cattle and not of the more productive exotic (i.e. grade) breeds.

Table 12: Impacts on number of animals (N=2558)

Dependent variable	Impact estimate (1)	Comparison Mean (2)	Treatment mean (3)
Number of cattle owned	***0.49	2.31	2.80
	(0.15)		
Number of exotic breed cattle owned	0.14	1.18	1.14
	(-0.28)		
Number of crossbred cattle owned	**0.29	0.68	0.97
	(0.14)		
Number of cows owned	***0.35	0.95	1.31
	(0.08)		
Number of milking cows owned	***0.44	0.57	1.01
	(0.07)		
Number of exotic breed cows owned	0.03	0.57	0.60
	(0.08)		
Number of crossbred cows owned	***0.16	0.28	0.44
	(0.06)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

We then investigated programme impacts on milk production, commercialisation and total value. Our findings indicated that the SDCP farmers were 8 percentage points more likely to sell milk in the morning and evening, although only the morning propensity to sell was statistically significant. The evidence suggests that the SDCP farmers were selling more milk in the market, although the results again were not statistically significant at the 10 per cent level of significance. Nevertheless, those farmers selling to the market could get a selling price that was 31 per cent higher than the selling price non-beneficiaries received. Overall, the total value of milk sold (calculated as the quantity of milk sold times the price) observed by the SDCP farmers, was 43 per cent higher than the value of the comparison group. Moreover, the SDCP farmers showed a higher level of milk production for animals at calving as well as higher total production, calculated as the sum of milk sold, consumed in the household, and lost.

Table 13: Impacts on milk production and total value (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Sold milk any time (yesterday)	**0.08	0.42	0.50
	(0.04)		
Sold milk in the morning (yesterday)	0.07	0.41	0.48
	(0.04)		
Sold milk in the evening (yesterday)	**0.06	0.20	0.26
	(0.03)		
Total litres of milk sold in the morning (yesterday)	0.12	0.86	0.98
	(0.10)		
Total litres of milk sold in the evening (yesterday)	0.11	0.38	0.49
	(0.07)		
Total litres of milk sold (yesterday)	0.18	0.96	1.14
	(0.11)		
Selling price of milk/litre (yesterday)	*0.31	1.81	2.13
	(0.18)		
Total value of milk sold	*0.43	2.47	2.90
	(0.25)		
Total milk production at calving	***0.58	1.33	1.92
	(0.12)		
Total milk production (yesterday)	***0.37	1.13	1.50
	(0.11)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$. All production values and selling prices are expressed in logs. Total milk production includes milk sold, consumed in the household, and lost.

These quantitative results were supported by qualitative data on the perceived effect of the SDCP on farmers' incomes from dairy farming. Without being prompted, FGD participants said they believed SDCP-sponsored activities had enabled them to more easily pay bills and send their children to school. One participant said, 'My milk production has increased and that's more income. With it, I have taken my children to school.' The perceived effect of the programme on income could partly result from the SDCP focus on recordkeeping. One farmer said, 'There is recordkeeping – tracking milk production and noting fall in its production – and getting to understand source of the problem and how to solve it.'

Although qualitative data cannot accurately indicate whether there was a real nominal increase in production and productivity per animal, FGD participants from treatment groups generally perceived a difference in production as a result of the SDCP. Multiple anecdotes indicated farmers' perceptions of increased productivity. One farmer described her increased yields: 'Initially I would only get three cups of milk, but currently my cow is producing seven bottles.' She also added that this increase helps to pay for school fees.

7.4.1 increased food security

FGD participants said they believed increased income as a result of the SDCP had enabled them to keep a variety of foods available in their households. One farmer said that his family's general health had improved, while others said they now consistently had tea with milk in their house. Farmers' perceptions of increased food security may partially have resulted from understanding better practices for growing crops, including the use of cow dung as fertiliser. One farmer said, 'The animal manure from both the goats and cattle is channelled on the farm, which in turn fastens the growth of crops and more yields – this too has ensured a consistent supply of food in key households.'

This evidence from the qualitative data is in line with the estimated impacts on the ability of the SDCP households to translate higher incomes into higher levels of food security. The survey includes details about food consumption within households to get a sense of how higher incomes have led to greater dietary diversity, which is a proxy for food security.

One such measure of food security was developed by the Food and Nutrition Technical Assistance Project (FANTA); greater values of the FANTA measure indicate more food insecurity. The FANTA project includes guidelines for a dietary diversity questionnaire that can be used at the household level. Specifically, this diversity measure involves calculating dietary diversity scores by summing up the number of food groups consumed by a household member over a reference period (our questionnaire set this to be the past seven days). Thus, the dietary diversity scores consist of a simple count of food groups that the household has consumed in the past seven days.

The results in Table 14 show that the SDCP households were more likely to have a more diverse food basket, especially of foods with higher levels of animal and vegetable proteins (e.g. red meats, milk products, and legumes such as beans, peas, lentils and nuts), and lower levels of tuber and fruit consumption, which are nonetheless still quite common among the SDCP farmers. Overall, the results provide evidence that programme beneficiaries exhibited higher levels of food diversification towards more nutritious food items.

Table 14: Impacts on food categories consumed in past seven days (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Cereals/grains	-0.00	0.99	0.98
	(0.01)		
Potatoes/yams/cassava	***-0.14	0.89	0.75
	(0.03)		
Vegetables	0.00	0.99	0.99
	(0.01)		
Fruits	***-0.11	0.89	0.78
	(0.03)		
Beans/peas/lentils/nuts	*0.03	0.88	0.91
	(0.02)		
Red meats/other organ meats	***0.17	0.36	0.54
	(0.04)		
Poultry	0.00	0.35	0.36
	(0.03)		
Eggs	0.02	0.60	0.62
	(0.04)		
Fresh fish/dried fish/shellfish	-0.02	0.37	0.35
	(0.03)		
Milk/cheese/yoghurt/other milk product	***0.09	0.80	0.89
	(0.03)		
Oils and fats	0.02	0.88	0.90
	(0.03)		
Sweets/sugar/honey	0.03	0.91	0.95
	(0.02)		
Other (condiments, coffee, tea)	**0.06	0.89	0.95
	(0.02)		
Household dietary diversity score	0.16	9.81	9.97
	(0.14)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

7.5 Sex differential effects

In this section, we investigate differential programme effects by sex of the household head. As indicated earlier, one of the goals of the SDCP was to serve women dairy farmers and the dairy groups they form. As a result, being a female-headed household increased the probability of participating in the programme by 11 percentage points (see Table 4). To the extent that female-headed households tend to have a lower socio-economic status than male-headed households (e.g. female-headed households have fewer working-age members, have fewer consumer durables or agricultural implements, and have fewer extensions of cultivated land), programme impacts may have been different in terms of this sex dimension.

We first explore programme impacts on household decision making. As part of the household survey, we collected information on numerous variables regarding who in the household participated in making decisions about dairy activities, including use of inputs and providers, management of money from milk sales, and requests for dairy-related trainings. All the questions considered allowed respondents to choose from five options about who made the decisions: household male, household female, joint household (male and female), non-household member and 'other'. For analytical purposes, we created indicator variables for all decision-making questions where the variable is equal to 1 if a female member made the decision, and 0 otherwise.

We present the results for some key variables in Table 15.[6] For most of the outcomes considered, the programme had a positive effect on the probability of making a decision relative to the comparison group. Treated households were 9 percentage points more likely to have a female managing the cash from fresh milk sold, relative to non-SDCP households. Also, the SDCP households were 7 and 11 percentage points more likely than the comparison group to have a female deciding the provider for bull and AI services, respectively. Lastly, household beneficiaries were more likely to have a female deciding the use of different services such as use of anthelmintics, tick control, vaccination and curative treatments.

[6] Note that for those decision-making variables with a low number of females making a given decision, the doubly robust estimator did not converge.

Table 15: Impacts on women's decision-making outcomes (N=2558)

Dependent variable	Impact estimate (1)	Comparison mean (2)	Treatment mean (3)
Female manages money from fresh milk sold	*0.09	0.21	0.30
	(0.05)		
Female decides own bull service provider	*0.07	0.11	0.18
	(0.04)		
Female decides AI service provider	***0.11	0.03	0.15
	(0.03)		
Female requested livestock best practices training	***0.11	0.10	0.20
	(0.03)		
Female requested cattle best practices training	0.06	0.17	0.23
	(0.04)		
Female decides use of anthelmintic	***0.27	0.17	0.45
	(0.05)		
Female decides use of tick control service	***0.19	0.21	0.40
	(0.06)		
Female decides use of vaccination service	***0.19	0.11	0.31
	(0.05)		
Female decides use of curative treatment service	*0.05	0.06	0.11
	(0.03)		

Notes: impacts were estimated using IPWRA (doubly-robust estimator). Robust standard errors are in parentheses. All estimations control for: household and household head characteristics; plot and crop production characteristics for primary and secondary rainy seasons; availability of other agricultural and livestock support programmes in the village; and pre-programme division-level characteristics, including milk density in litres per km^2, proportion of people below the poverty line, and travel time to the nearest cooling and urban centre. *** $p < 0.01$; *** $p < 0.05$; *** $p < 0.1$

The overall results indicate that the SDCP programme had a positive impact on the welfare of female dairy farmers by empowering women to make decisions related to dairy production. These quantitative results are encouraging, since the qualitative results did not provide insight into the gender dynamics of dairy farming. Gender was included as a topic in FGDs, but farmers had very little interesting information to share about their perspectives on gender.

7.6 Cost analysis

The total cost of the SDCP was KSH 1,515,614,910 for the period from 2006 to 2016. Appendix Table B1 includes the yearly expenses across categories. The largest category of expenses was 'technical assistance, training and workshops', which amounted to 33.4

per cent of the overall budget. This large share of the budget is not surprising since many of the programme activities related to this line item. Other higher categories of expenses included ‘salaries and allowances’ (17.5 per cent), ‘vehicle and office operating costs’ (13.2 per cent) and ‘vehicles, equipment and materials’ (14.5 per cent).

To compare the costs to the benefit that each farmer received from the programme, we first need an estimate of the costs per beneficiary. To determine the number of beneficiaries, we examined the dairy groups inventory from the programme counties and divisions, summarised in Appendix Table B2. Most of the inventories were updated in 2012, with some having been revised in 2011. Across treatment divisions, there were 505 dairy groups, with 15,535 total dairy group members. On average, each dairy group had 30.8 members. Taking the total costs from Appendix Table B1 and the total number of beneficiaries from Appendix Table B2, we can estimate that the total costs per beneficiary were around KSH 97,561 (equal to KSH 1,515,614,910 divided by 15,535).

We next need a measure of the benefits of the programme at farmer level. While we recognise that the programme is associated with a variety of farmer-level benefits, we focus on the benefit of increased milk production for the purpose of this analysis, since that was one of the main goals of the SDCP. The unconditional average amount of milk produced the preceding day for a farmer in the comparison group was 3.81 litres, where this estimate includes zero values for comparison group farmers who produced no milk the preceding day.

Our impact estimates (Table 13) suggest that the programme increased milk production by 37 per cent, suggesting that the additional amount of milk produced yesterday by a farmer in the treatment group is 1.41 litres. Because our estimate of the milk produced the preceding day included values of zero for when the farmer did not produce milk that day, these estimates reflect the amount produced on an average day. We assume this average level of production throughout the year and multiply the estimate of 1.41 additional litres per day by 365 days in a year to estimate that the additional milk produced by a farmer in a year is 514.65 litres. In monetary terms, this is equivalent to a benefit of KSH 20,586 per farmer since the unconditional mean of the price per litre of milk, which excludes zeros for those who did not sell milk, is KSH 40. This benefit estimate corresponds to the 2016 timeframe, as the data on the preceding day’s milk production was collected in 2016.

With an estimated total cost per beneficiary of KSH 97,561 and an estimated per farmer benefit of KSH 20,586 in the year 2016, we calculate that it would take around 4.74 years (equal to KSH 97,561 divided by KSH 20,586) to break even; that is, for the benefits to equal the costs, assuming the benefits of KSH 20,586 remain constant across years. This estimate of the number of years to break even seems reasonable, because although there were not likely to have been any benefits at the beginning of the programme (when upfront costs were associated with setting up the SDCP), farmers were likely to have begun to benefit from the programme before 2016.

Furthermore, to the extent that farmers continue to employ the best practices the programme advocated, the benefits may extend into future years. In addition, these estimates seem reasonable because we only examined the benefits of increased milk production, and the SDCP benefited treatment farmers in a variety of dimensions.

8. Discussion and conclusions

Overall, the results suggest that the SDCP was successful in increasing milk production, but with more limited – though positive – impacts on increasing milk marketing and increasing milk prices received by smallholders. Thus, we first summarise results in milk production then on milk marketing. We end the section with a summary of the key challenges remaining.

8.1 Milk production

The SDCP households surveyed were more likely to have received information on all of the practices the SDCP promoted, versus control households; they were also more likely to have adopted those practices. In addition, the SDCP households were more likely to have crossbred cows, used AI services and accessed a wider range of health services. Finally, they were also more likely to have adopted recommended management practices and investments, including practising zero grazing, having concrete floors and feeding concentrates. In fact, control households did not perform better on any measure of input use, management or investment, with just one exception: the SDCP households were less likely to clean cows' teats before milking. Overall, these improved input and management practices led to increased milk production.

In addition to increasing milk production per cow, another key objective of the SDCP was to even out milk production across the year by increasing access to adequate feed and fodder throughout the year. Because of the difficulty in obtaining recall data on milk production throughout the year, we have no direct evidence on whether SDCP farmers actually have more even outputs. We found that the SDCP farmers were more likely to adopt fodder establishment, hay making, silage making and conservation of crop residues, all of which should increase availability of adequate fodder throughout the year. However, adoption rates for each of these practices were below 15 per cent, suggesting there may still be room to improve activities and trainings aimed at evening out fodder availability, and thus milk production, over the year.

Results from the household survey were consistent with results from the dairy group survey and the FGDs. The SDCP dairy group respondents noted increased access to trainings on almost all topics. In particular, 79 per cent of the SDCP dairy groups noted that at least one training had been provided on fodder management, well above the 36 per cent in control dairy groups. However, household respondents were less likely to receive information on various practices, implying that relatively few households in the dairy groups attended these trainings. In fact, the number of dairy group leaders and members attending trainings was somewhat limited, at about 16 attendees for the SDCP trainings and slightly fewer in control groups. This implies scope for increasing dissemination of knowledge learned at trainings to the wider group of dairy farmers who could not attend.

FGD participants thought that the SDCP project helped them achieve higher milk yields, higher income and greater food security. The most frequently discussed benefit was improved animal management. Many participants also discussed the need to make and preserve fodder to ensure year-round availability of fodder for their dairy cows, and the SDCP farmers appear to have a more sophisticated knowledge of fodder management.

However, they also expressed that some farmers faced barriers in adopting best practices. In part, FGD participants thought the increased availability and consistency of technical services the SDCP provided would increase the number of adopters. Additionally, many farmers had misconceptions about how AI works, and many also thought the price was too high, particularly given that success first time is not guaranteed.

8.2 Milk and input markets

Total milk production and milk sold were higher for the SDCP, although the impact was higher for production than for milk marketed. In part, this appears to have been because the SDCP producers were more likely to have sold any milk the preceding day; in particular, they were more likely to have sold milk in the evening, and in both the morning and the evening. This evidence suggests that there were positive impacts on milk marketing.

On the other hand, from the household survey, the percentage of households receiving information on market-related topics was generally lower than for production-related and farm management topics for both the SDCP and control groups, although the SDCP households were more likely to have received market-related information than controls. Just nine per cent of households noted that they had received information on fresh milk marketing, with even fewer having received information on value addition and cooperative milk marketing. Similarly, in the dairy group surveys, 59 per cent reported trainings related to marketing, which is significantly below the number reporting trainings on production-related activities such as fodder management (79 per cent).

The dairy group survey also reflects more limited impacts on marketing services they provided to their members. First, there were limited differences between treated and control dairy groups in the types of services offered. Only 20 per cent of the SDCP groups facilitated links between members and input suppliers, while 24 per cent facilitated linking members to milk purchasers, which is similar to the percentages observed in control groups. Just 9 per cent of the SDCP dairy groups contracted with input suppliers on behalf of members, and 11 per cent contracted with milk purchasers on behalf of members. However, while 11 per cent was fairly low, it was significantly higher than control groups that contracted with milk purchasers, at 2 per cent.

Some treatment FGDs said they thought the project had reduced transaction costs, and one participant said that it had become easier to find a steadier customer base. On the other hand, many participants also noted that entering the market was still a challenge for many farmers, and the FGD interviewers noted that farmers expressed a lack of understanding of how to expand their business.

One of the more successful initiatives to help farmers access input and output markets appears to have been the expansion of access to credit, as primarily documented in the dairy group surveys and FGDs. The SDCP dairy groups were more likely to have accessed a wider range of sources of finance, including micro-finance and commercial sources, but particularly local savings and loan clubs. Similarly, treatment farmers in FGDs frequently mentioned the SDCP trainings, which led to the establishment of merry-go-round and table banking groups.

8.3 Challenges

Overall, the SDCP activities contributed to a positive overall picture of the impact and perceptions of the SDCP on Kenyan dairy farmers, particularly for dairy production. However, insights from the qualitative data also point to challenges in service provision, use of technical knowledge and skills, and institutional coordination that could have enhanced the programme.

8.4 Uptake and ongoing services

Key informants notably expressed different primary challenges to dairy farming commercialisation in Kenya than their beneficiary counterparts, albeit that the challenges were also elements that the SDCP actively aims to address. In addition to mentioning challenges associated with increasing uptake of improved feed management, marketing and dairy farm management, key informants' top two most frequently mentioned challenges were: (1) the availability and consistency of technical services that relate to the SDCP; and (2) the lack of uptake and use of available knowledge and technical skills among farmers.

Discussion of these two primary challenges was related. Many key informants particularly mentioned **extension among government-provided services that suffer because of lack of staff and funding**. One informant said, 'All development projects within the county go through the same officers within the county, [and the] number of staff are not enough to reach out to those communities.' This statement highlights that despite the provision of SDCP activities, a staffing bottleneck may prevent these services from being implemented as intended. One trainer from the Ministry of Agriculture recommended that they 'visit the groups mid-year to see if they implement what we have trained them on and if they have any technical challenges.'

Given that such services aim to help farmers with technical knowledge, it makes sense that the **use of available knowledge and technical skills** was also a notable challenge. Even in cases where key informants believed that information about technical skills in farm management had reached the farmers, they then mentioned challenges with farmers being able to use the knowledge because of capital constraints. One key informant pointed out this constraint: 'Much as we offer the extension, services may not be adequate and…farmers are slow in taking up the services that we offer.' The same key informant connected the challenge of services with financial challenges on the part of the farmer: 'Like in AI – as much it is subsidised, farmers still complain that the price is too high. The farmers have not fully the idea of commercialisation.'

In addition to lack of capital to consistently implement farm and animal management practices, farmers also face challenges **negotiating their entrance into markets**. One dairy group chair said of the SDCP, 'Now that they have improved our lives by providing machines for handling milk, they should try to continue supporting us to see that we can even package and market for ourselves.' This indicates that despite having had trainings on marketing, not all farmers have put the changes into practice.

Farmers were also unable to concretely identify how to go about expanding their business in other ways. One dairy group member said, 'I would like if we could get capital to invest in our business of making feeds and we market them and sale to other

places.' However, capital was an issue in this group, as well as others. Another farmer said, 'If we can get our own doctors and at least our group to pay for the treatment then they can deduct the cost in milk, then we would really prosper.'

8.5 Institutional coordination and linkages

FGDs indicated there were **challenges with linking farmers and dairy groups to various service providers**, and that farmers did not always have the knowledge to negotiate to their benefit. Key informants agreed that bringing service providers together continued to be a challenge. One key informant said there continued to be 'inadequate coordination of the value chain actors, especially in terms of contract terms and contractual agreements.' Farmers generally did not discuss interactions with value chain actors other than AI providers and veterinary officers. One farmer said:

> [With] the service providers, it's on need basis; while the agricultural extension officers, we meet them on weekly basis as they liaise with other stakeholders, like [the] World Bank [or others who] deal with breeding of animals and rearing of chickens.

This indicated that, in some cases, they relied on extension officers for coordination of major animal management activities.

There was **no mention among dairy group farmers or their chairs of any formal contractual arrangements with banks, service providers or other actors in the value chain**. The lack of agreements was evident when one key informant said he wished there could be 'proper structures on marketing of milk and increased production of milk from individual farmers. We should not experience [the] case that happened in Nyandarua where they poured milk because they did not have a place to take it.' One farmer explained that his group had entered into a marketing agreement:

> We have approached an organisation called [the] Inter Region Economic Network and they came in and researched and decided they will bring us marketers for our products, and as a result we shall sell our milk as individuals. So the agreement was they sell the milk in very huge amounts and they keep 15 per cent and us the rest 85 per cent.

However, there was no other mention of such arrangements.

8.6 Funds disbursement

Finally, a major challenge to programme implementation that multiple key informants cited was the **inability to get SDCP funds through the various levels of the government to enable timely implementation of initiatives on the ground**. Key informants mentioned particular elements of funding, including disbursing funds from the national level of the government, enabling payments to service providers, and – even among farmers – the inability to purchase services, which some informants speculated came from an impression among farmers that they should not have to pay for services or should receive a payment for attending trainings. This challenge among farmers aligned with the existing idea that farmers had not fully taken up the trainings. One informant reiterated, 'The challenges on finances…hamper implementation, [including with] group

issues on governance, leadership and resource mobilisation. They may have the knowledge, but still those conflicts arise.'

At the level of government implementation, one key informant said there was a challenge with 'authority to incur expenditure; you cannot actually be allowed to spend before it's signed.' At the time of the interview, it had been more than a month since the office was supposed to have received the authority to incur expenditure, but it still had not received it. Key informants provided mixed responses on whether devolution caused any of these challenges with implementation of the SDCP, although some respondents thought having an extra layer of bureaucracy might contribute to the slowdown of funds and information flow (despite the programme still being implemented from the national level).

8.7 Internal and external validity

The results of the evaluation of the SDCP programme can be used to inform the design and implementation of similar dairy policies and projects elsewhere in Kenya and in other developing countries. First, there are at least 16 counties in Kenya that, although not served by the programme, have comparably resource-poor dairy farmers who we expect would benefit in similar ways from SDCP activities should they receive the programme.

Although road densities might differ from the densities that IFAD and the GoK selected, the selected districts had high road densities and were also relatively contiguous in order to ease SDCP management. Thus, the results should extend to similar areas in Kenya with high road densities that are not necessarily contiguous.

Second, dairy farmers across East Africa share many of the features and challenges that smallholder dairy farmers face in Kenya. Most small dairy farmers from neighbouring countries live in subtropical areas suitable for dairy cattle and produce milk as an important part of their diet (Thorpe et al. 2000). These dairy farmers also face similar challenges, such as lack of training and availability of production inputs that prevent them from commercialising their dairy products in the most efficient way.

More generally, the effects of the SDCP programme can provide insights about other types of agricultural extension programmes to smallholder farmers in developing countries. A key objective of agricultural extension is to increase farmers' knowledge about agricultural practices, which in turn should have an effect on productivity. The traditional extension model, known as a 'training and visit' extension, is characterised by government-employed extension agents visiting farmers individually or in groups to demonstrate agricultural best practices (Anderson and Birner 2007).

However, the training and visit approach has been criticised for several reasons. First, given limited transportation infrastructure in rural areas and high costs of delivering information in person, in some developing countries the reach of extension programmes is limited. Second, as it is costly to provide agricultural extension programmes to farmers on a recurring basis, infrequent and irregular meetings limit the ability to provide timely information, which further limits the ability of farmers to provide timely information. Last, the information provided to farmers is usually too top-down, which results in inadequate diagnosis of the difficulties farmers currently face, as well as information that is often too technical for semi-literate farming populations.

The SDCP programme has key features that may help overcome some of the limitations of this traditional extension approach. In particular, a key aspect of the programme is that it aims to build the capacity of groups of farmers through trainings that allow them to identify their most pressing needs and design work plans to request more targeted assistance from the programme in terms of information and resources. By increasing the capacity of groups and not only individual farmers, the programme may not only be more cost efficient, but also generate positive externalities for other farmers in the area who may later become part of the groups.

The programme incentivises farmers to organise and compete for additional funds to improve their productivity. Our results are consistent with this approach to extension; however, our results also suggest that more needs to be done to help group members who participate in trainings to subsequently share the knowledge they learn with group members who did not participate.

We also believe that the quantitative results provide useful information on the impacts of a complex project for smallholder dairy farmers in other regions in Kenya and in neighbouring countries who face similar market characteristics. However, given the specificities associated with milk production and marketing, we do not expect that the results would be easily generalisable to other cash crop markets.

Online appendixes

Note to the reader: This appendix is only available online and has been published as it was received from the authors. It has not been copy-edited or proofed by 3ie and can be accessed from below links.

Online appendix A: Sample size and power calculations

http://www.3ieimpact.org/sites/default/files/2019-01/tw4102-ie73-kenya-dairy-commercilisation-appendix-a.pdf

Online appendix B: Cost data for the programme

http://www.3ieimpact.org/sites/default/files/2019-01/tw4102-ie73-kenya-dairy-commercilisation-appendix-b.pdf

References

Anderson, J and Birner, R, 2007. *How to make agricultural extension demand-driven? The case of India's Agricultural Extension Policy*. IFPRI Discussion Paper 00729. Washington, DC: International Food Policy Research Institute.

Atieno, R and Kanyinga, K, 2008. *The revitalization of Kenya Cooperative Creameries: the politics of policy reforms in the dairy sector in Kenya*. Future Agricultures Research Paper 009. Future Agricultures Consortium.

Bernard, HR, 2011. *Research methods in anthropology: Qualitative and quantitative approaches*. 5th ed. Lanham, MD: AltaMira Press.

Burke, WJ, Jayne, T, Freeman, H and Kristjansen, P, 2007. *Factors associated with rural farm households' movements into and out of poverty: the rising importance of livestock*. International Development Working Paper 90. Department of Agricultural Economics: Michigan State University.

Burke, W, Myers, M and Jayne, T, 2015. A triple-hurdle model of production and market participation in Kenya's dairy market. *American Journal of Agricultural Economics,* 97(4), pp.1227–1246.

FAO, 2017. *Global Information and Early Warning System Country Brief on Kenya*. 25 April. Available at: <http://www.fao.org/giews/countrybrief/country.jsp?code=KEN> [Accessed 16 June 2017].

Gelan, A and Muriithi, B, 2015. Examining returns to scale in smallholder dairy farms in East Africa. *Quarterly Journal of International Agriculture*, 54(3), pp.237–259.

Glaser, B and Strauss, A, 2009. *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Alpine.

Heckman, J, Ichimura, H, Smith, J and Todd, P, 1998. Characterizing selection bias using experimental data. *Econometrica,* 66(5), pp.1017–1098.

Heckman, J, Ichimura, H and Todd, P, 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies,* 64(4), pp.605–654.

Herrero, M, Thornton, P, Bernues, A, Baltenweck, I, Vervoort, J, van de Steeg, J, Makokha, S, van Wijk, M, Karanja, S, Rufino, M and Staal, S, 2014. Exploring future changes in smallholder farming systems by linking socio-economic scenarios with regional and household models. *Global Environmental Change*, 24, pp.165–182.

Holloway, G, Nicholson, C, Delgado, C, Stall, S and Ehui, S, 2000. Agroindustrialization through institutional innovation: transaction costs, cooperatives and milk-market development in the east-African highlands. *Agricultural Economics*, 23(3), pp. 279–288.

Imbens, G and Wooldridge, J, 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature,* 47(1), pp. 5–86.

McKenzie, D, 2011. Power calculations for propensity score matching. 20 November 2011. *Development Impact Blog*. Available at: <http://blogs.worldbank.org/impactevaluations/power-calculations-for-propensity-score-matching> [Accessed 5 March 2017].

Muriuki, HG, 2003. *A review of the small scale dairy sector – Kenya*. Rome, Italy: Food and Agriculture Organisation. Available at: <http://www.fao.org/fileadmin/templates/ags/docs/dairy/P1assessmentkenya.pdf> [Accessed 20 April 2017].

Nafula, S, 2013. The impact of enhanced nutritional and feeding practices on growth and health of dairy calves in Mukurweini District of Kenya. Masters of Veterinary Medicine thesis, University of Nairobi. Available at: <http://erepository.uonbi.ac.ke/bitstream/handle/11295/72378/The%20Impact%20Of%20Enhanced%20Nutritional%20And%20Feeding%20Practices%20On%20Growth%20And%20Health%20Of%20Dairy%20Calves%20In%20Mukurweini%20District%20Of%20Kenya.pdf?sequence=3&isAllowed=y> [Accessed 12 May 2017].

Ouma, R, Njoroge, L, Romney, D, Ochungo, P, Staal, S and Baltenweck, I, 2007. *Targeting dairy interventions in Kenya: A guide for development planners, researchers and extension workers*. ILRI Manuals and Guides No. 1. Nairobi, Kenya: International Livestock Research Institute.

Robins, J and Rotnitzky, A, 1995. Semiparametric Efficiency in Multivariate Regression Models with Missing Data. *Journal of the American Statistical Association*, 90(429), pp.122–29.

Robins, J, Rotnitzky, A and Zhao, L, 1995. Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association*, 90(429), pp.106–21.

Sacerdote, B, 2004. Fixing broken experiments using the propensity score. In: A Gelman and X-L Meng, eds. 2004. *Applied Bayesian modeling and causal inference from incomplete-data perspectives.* Hoboken, NJ: Wiley, pp. 61–71.

Staal, S, Delgado, C and Nicholson, C, 1997. Smallholder dairying under transactions costs in East Africa. *World Development*, 25, pp.779–794.

Thorpe, W, Muriuki, HG, Omore, A, Owango, MO and Staal, S, 2000. *Development of smallholder dairying in Eastern Africa with particular reference to Kenya*. Paper prepared for the UZ/RVAU/DIAS/DANIDA-ENRECA Project Review Workshop. Harare, Zimbabwe.

Van der Laan, M and Robins, J, 2003. *Unified Methods for Censored Longitudinal Data and Causality*. New York: Springer, Physica-Verlag.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/evidence-hub/impact-evaluation-repository

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, *3ie Impact Evaluation Report 67.* Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, *3ie Impact Evaluation Report 63.* Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, *3ie Impact Evaluation Report 58*. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional status, *3ie Impact Evaluation Report 55.* Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

Evaluating the effectiveness of computers as tutors in China, 3ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, *3ie Impact Evaluation Report 34.* Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, *3ie Impact Evaluation Report 24.* Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, *3ie Impact Evaluation Report 21.* Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, *3ie Impact Evaluation Report 20.* Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, *3ie Impact Evaluation Report 17.* Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9.* Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8.* Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5*. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2.* Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1.* Martinez, S, Naudeau, S and Pereira, V, 2012.

This evaluation examines the impacts of the Smallholder Dairy Commercialization Programme, which was designed to enhance dairy farming productivity, reduce production costs and to strengthen market linkages. The findings from this impact evaluation suggest that the programme was successful in increasing milk production, but had limited impacts on increasing the quantity of milk sold or the prices received by farmers.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org