DISCUSSION PAPER SERIES

DP15220

SOCIAL LEARNING IN AGRICULTURE: DOES SMALLHOLDER HETEROGENEITY IMPEDE TECHNOLOGY DIFFUSION IN SUB-SAHARAN AFRICA?

Luc Behaghel, Jeremie Gignoux and Karen Macours

DEVELOPMENT ECONOMICS

ISSN 0265-8003

SOCIAL LEARNING IN AGRICULTURE: DOES SMALLHOLDER HETEROGENEITY IMPEDE TECHNOLOGY DIFFUSION IN SUB-SAHARAN AFRICA?

Luc Behaghel, Jeremie Gignoux and Karen Macours

Discussion Paper DP15220
Published 28 August 2020
Submitted 09 August 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SOCIAL LEARNING IN AGRICULTURE: DOES SMALLHOLDER HETEROGENEITY IMPEDE TECHNOLOGY DIFFUSION IN SUB-SAHARAN AFRICA?

Abstract

Evaluating a large-scale program for dairy farmers in Uganda, we show that a simple version of the "contact farmer" extension model can meaningfully increase smallholder farmers' revenues. While the program provides no monetary incentives, we find evidence that two other ingredients – backstopping by professional extension agent and advertising pro-social motivation – reinforce its impacts. Though it has been hypothesized to be a major impediment to social learning in Sub-Saharan African agriculture, we do not find smallholder heterogeneity to condition the effectiveness of the approach: farmer trainers trained to take this heterogeneity into consideration do not perform better; moreover, we find no statistical evidence that program effects vary by farmers' characteristics.

JEL Classification: O12, O13, O33, Q16

Keywords: Agricultural Productivity, Heterogeneity, extension, Livestock, Social learning

Luc Behaghel - luc.behaghel@psemail.eu
Paris School of Economics and INRAE

Jeremie Gignoux - jeremie.gignoux@psemail.eu
Paris School of Economics and INRAE

Karen Macours - karen.macours@psemail.eu
Paris School of Economics and INRA and CEPR

Acknowledgements
We thank the International Initiative for Impact Evaluation for funding the main body of research reported in this study. We also thank the CGIAR Research Program on Policies, Institutions and Markets and ISPC-SPIA under the grant "Strengthening Impact Assessment in the CGIAR System (SIAC)" for supplemental funding. We are grateful to the World Agroforestry Centre (ICRAF) and the East African Dairy Development Project, led by Heifer International, for hosting the research at their project sites in Uganda, and to Filder Aryemo and the Gaplink field team for excellent data collection. We thank Steve Franzel (ICRAF) for his constant and insightful support. Finally, we thank the dairy producer organizations and farmers who participated in the study for giving us so much of their time and providing us with the needed information. The research team on this project involved Olivia Bertelli, Monica Karuhanga Beraho, Jane Kugonza, Margaret Najjingo Mangheni, and Judith Oduol. Cansin Arslan, Josué Awonon, Caroline Liqui Lung, Sebastián Sandoval Olascoaga, Ibrahim Ouattara, Victor Saldarriaga, and Mélanie Staquet provided invaluable research assistance. We thank Denis Cogneau, Marc Gurgand, Yagan Hazard, Alain de Janvry, Sylvie Lambert, and seminar participants at PSE, the World Agroforestry Congress (2019), and CIRAD-INRAE Montpellier. All remaining errors are ours.

Social learning in agriculture: does smallholder heterogeneity impede technology diffusion in Sub-Saharan Africa? *

Luc Behaghel† Jérémie Gignoux ‡ Karen Macours§

August 9, 2020

Abstract

Evaluating a large-scale program for dairy farmers in Uganda, we show that a simple version of the "contact farmer" extension model can meaningfully increase smallholder farmers' revenues. While the program provides no monetary incentives, we find evidence that two other ingredients – backstopping by professional extension agent and advertising pro-social motivation – reinforce its impacts.

Though it has been hypothesized to be a major impediment to social learning in Sub-Saharan African agriculture, we do not find smallholder heterogeneity to condition the effectiveness of the approach: farmer trainers trained to take this heterogeneity into consideration do not perform better; moreover, we find no statistical evidence that program effects vary by farmers' characteristics.

*We thank the International Initiative for Impact Evaluation for funding the main body of research reported in this study. We also thank the CGIAR Research Program on Policies, Institutions and Markets and ISPC-SPIA under the grant "Strengthening Impact Assessment in the CGIAR System (SIAC)" for supplemental funding. We are grateful to the World Agroforestry Centre (ICRAF) and the East African Dairy Development Project, led by Heifer International, for hosting the research at their project sites in Uganda, and to Filder Aryemo and the Gaplink field team for excellent data collection. We thank Steve Franzel (ICRAF) for his constant and insightful support. Finally, we thank the dairy producer organizations and farmers who participated in the study for giving us so much of their time and providing us with the needed information. The research team on this project involved Olivia Bertelli, Monica Karuhanga Beraho, Jane Kugonza, Margaret Najjingo Mangheni, and Judith Oduol. Cansin Arslan, Josué Awonon, Caroline Liqui Lung, Sebastián Sandoval Olascoaga, Ibrahim Ouattara, Victor Saldarriaga, and Mélanie Staquet provided invaluable research assistance. We thank Denis Cogneau, Marc Gurgand, Yagan Hazard, Alain de Janvry, Sylvie Lambert, and seminar participants at PSE, the World Agroforestry Congress (2019), and CIRAD-INRAE Montpellier. All remaining errors are ours.

†Paris School of Economics, INRAE, luc.behaghel@psemail.eu

‡Paris School of Economics, INRAE, jeremie.gignoux@psemail.eu

§Paris School of Economics, INRAE, karen.macours@psemail.eu

1 Introduction

Agricultural productivity in Sub-Saharan Africa is lagging behind, and recent effort to close the gap has had modest aggregate effect at best (Goyal & Nash 2017). Productivity growth, the predominant source of agricultural growth in the rest of the developing world, has remained extremely weak in most Sub-Saharan countries. A vast economic literature has considered possible explanations of, and solutions to, this agricultural productivity gap. It has shown the limited role of some factors (such as pure credit constraints) and stressed the importance of others (such as access to input and output markets)(de Janvry et al. 2017). And while this literature points to the lack of access to information on the profitability and the optimal use of technologies as a potential important barrier to technology adoption, many questions remain on how best to overcome the information barriers for smallholder farmers in Sub-Saharan Africa (de Janvry & Sadoulet 2019).

This paper evaluates the impact of a program in which the World Agroforestry Center (ICRAF) trains volunteer "Farmer Trainers" (FTs) on multiple strategies to improve feeding of their dairy cows; FTs in turn train other farmers in their villages. The evaluation is based on a clustered randomized control trial (RCT) taking advantage of the roll-out of the FT program to the Busoga region of Uganda. The impact of the FT program is measured after two years, by comparing 51 treatment areas (consisting of 411 villages) that received FTs in year 1, to 26 control areas (216 villages) that received them only in year 3. Among treated villages, we also test the impact of three variations aimed to increase the program's effectiveness, by providing (cross-randomized) subsets of FTs with (i) vouchers that cover the cost of visits by professional extension agents to the FT ("linkage" variation); (ii) extra training to adjust their teaching to specific needs of fellow farmers ("needs assessment" variation); (iii) a metallic signpost to put in front of their compound indicating that they are FTs and tracking their activity ("signpost" variation). The variations are randomized at the village (FT) level.

The overall program has positive impacts all along the hypothesized causal chain, from knowledge to technology adoption, production, and profits. Impacts on FTs' knowledge are large (+48% standardized effect), and are transmitted to other farmers in their village (+8% standardized effect). FTs and other farmers report increased use of feeds and feeding practices, and planting more feeds on their plots. They report higher milk production: among farmers, production increases by about 13% of a standard deviation. Increases are concentrated on crossbred animals. We corroborate the milk increases through surprise visits during milking sessions at a subsample of farms; we also observe significant increases in cows' weight, consistent with better feeding. Last, though impacts on dairy profits are not precisely estimated, they tend to be positive.

Our first contribution is to provide large-scale evidence on the effectiveness of a contact farmer extension model while also providing experimental evidence on the mechanisms contributing to this effectiveness. The contact farmer (or farmer-to-farmer) model is extremely common across the developing world. The version of the model we evaluate was designed to have a relatively low cost per dairy farmer reached with three main features: (i) a participatory process to select one farmer trainer per village; (ii) a moderately intense training provided by professional staff to groups of about 40 FTs (3 days in the first year, 2 days in the second one); (iii) the absence of direct monetary incentives.

Finding statistically significant and economically meaningful impact is important, and goes against the long-standing pessimistic view on the feasibility of cost-effective extension (Anderson & Feder 2007). The difficulty of monitoring and incentivizing agricultural extension workers, and the reality of high farmer-extension agent ratios, in particular are often thought to be major constraints. A recent literature studies how to circumvent some of these challenges through modifications in the contact-farmer extension models, focusing on the role of incentives, and in particular on the selection (identity) of the contact farmers and their ability to reach large parts of village networks (BenYishay & Mobarak 2018, Beaman et al. 2018, BenYishay et al. 2020, Emerick 2018, Beaman & Dillon 2018). As summarized by Magruder (2018), this literature shows that typically strong information frictions remain in these approaches, even if (some) social learning occurs.[1]

Our second contribution relates to the running hypothesis that the heterogeneity of farmers and farm conditions can significantly impede agricultural technology diffusion through social learning in Sub-Saharan Africa. Indeed, a possible advantage of the contact-farmer model that has received less attention is its ability to draw on local knowledge regarding conditions and constraints to target extension messages. Such customization is believed to be particularly valuable given evidence on the large heterogeneity in returns to improved technologies (Marenya & Barrett 2009, Suri 2011).[2] But heterogeneity of response to inputs within a village can also complicate the social learning process (Munshi 2004). Large within-village heterogeneity in soil characteristics in particular has been linked to slower diffusion (Tjernström 2017, Assunção et al. 2019), and may mean that village-level interventions fail to circumvent the key heterogeneity. One recent study that specifically analyzes this question experimentally for fertilizer recommendations finds, however, that individual customization did not lead to higher take up of fertilizer recommendation than village level customization (Corral et al. 2020).

Whether and how the contact farmer model is suited to address the relevant heterogeneity and customize messages at the right level is therefore an open question. The selection and

[1] Hanna et al. (2014) further show that learning from demonstration trials can be imperfect because farmers may fail to notice important features.

[2] Marenya & Barrett (2009) show that yield response to mineral fertilizer in Kenya is very heterogeneous, leading to negative profitability for the one-third of farmers with the most deficient soils. Related, Suri (2011) shows that farmers with the highest estimated gross returns to hybrid seeds in Kenya do not use it, but that their returns are correlated with high costs of acquiring the technology, while other farmers with lower returns are adopting. This can help explain evidence regarding limited social learning (Duflo 2018).

training of one FT with deep knowledge of local conditions in each village is a first potential mechanism to target village-level heterogeneity. The focus of the trainings on conveying a set of key principles, rather than focusing uniquely on specific practices (Stevenson et al. (2019)), may also have facilitated FTs ability to adapt to those local conditions. The "needs assessment" variation in our experiment was designed to test whether enhancing FTs ability to target individual heterogeneity, by providing tools to the FT in order to adapt his teaching to the different farmers, improves outcomes. We do not find this variation to have any effect, which may be due to the absence of significant individual heterogeneity in extension needs in our context, or to the limitations of our low-tech and low-cost approach to customization. The ex post analysis of heterogeneous treatment effects, along the lines set by our pre-analysis plan (based upon theoretical priors and qualitative evidence on relevant dimensions of heterogeneity) as well as through a machine-learning approach following Chernozhukov et al. (2018) also does not show significant heterogeneity. On the other hand, intra-village correlations in baseline knowledge and practices are high, and access to FTs (knowledge) appears relatively equitable within villages. Hence even without individual customization, topics selected by the FT for training could have relevance for most farmers in a village. Together, these results suggest that individual heterogeneity in farmers' circumstances, if it exists, may not be obvious to elicit and respond to, but the local FTs system may well be effective at accounting for heterogeneity across villages.[3]

Finally, this paper also adds to the literature highlighting the potential of livestock interventions in developing countries. The evaluations of the ultra-poor programs (Banerjee et al. 2015, Bandiera et al. 2017) shown that multi-faceted programs involving livestock asset transfers, training and coaching can successfully change the long-term trajectory of the poorest. A review of wider (mostly observational) literature further suggest the promise of dairy development projects for reducing poverty and improving nutrition (FAO et al. 2018). Interventions that help shift farmers to more efficient feeding practices are hypothesized, through their impact on productivity, to help decrease the environmental footprint of livestock farming by (Havlik et al. 2014). This paper contributes with experimental evidence of the causal chain from knowledge and improvements in feeding practices to milk yield and profits of an extension intervention implemented at scale.

The next section presents the farmer trainer program and its expected effects. Section 3 details the impact evaluation design and the data sources. The effects of the FT program and their underlying mechanisms are studied in Section 4, while Section 5 returns to the heterogeneity discussion. Section 6 provides concluding comments. Supplementary materials including the pre-analysis plan, power computations, and survey instruments are available in

[3]As such our study also relates to recent work on customization of agricultural advise through IT methods. Recent efforts in particular try to us IT methods to incorporate information about local conditions and send targeted messages regarding specific practices adapted to local conditions (Aker 2011, Cole & Fernando 2020, Fabregas et al. 2018, Casaburi et al. 2019).

a web appendix.[4]

2 The Farmer Trainer program

The farmer trainer (FT) program is a component of a broader program aiming to develop the dairy sector, the East Africa Dairy Development Project (EADD), led by a consortium of NGOs and international agricultural research centers.[5] The FTs, who are volunteers, are trained in different improved feed practices. They are subsequently asked to disseminate this information through demonstrations, teaching, and, when possible, access to seeds/planting material. The FT program is expected to complement the other components of EADD, which aim at empowering dairy producer organizations to develop new marketing channels for the milk (including through the investment in infrastructure such as cooling tanks), to promote the diffusion of crossbred cattle (in particular through artificial insemination), and to provide dairy farmers with a variety of veterinary and extension services. Improved feeding practices serve the overarching goal of developing a sustainable dairy sector that brings dairy farmers a steady source of income. Importantly, the genetic improvement (cross breeding) and improved feeding practices are considered mutually reinforcing for productivity gains. FTs are viewed by EADD as the front-line of extension departments set up by the producer organizations, who through their local presence are well placed to foster that change.

The FT program involves several steps. FTs receive an initial two-day training in feeding practices by the ICRAF program staff, which provides them with technical and practical knowledge. The training is focused on conveying key principles for dairy diets (balanced nutrient content, smoothing seasonal fluctuations, water access, etc) and includes showcasing practical techniques and limited distribution of pasture seeds and seedlings. FTs then receive "refresher trainings" at a biannual frequency. This first step is meant to increase their knowledge of feeds and to lead them to implement new practices in their farms. FTs are also given advice on how to train other farmers, and asked to do so. The second step is therefore to have them organize training sessions in their village, and to visit fellow farmers to provide them with advice. The expected result of this second step is knowledge and increased adoption of better feeding practices by farmers. Adoption may involve planting grasses or fodder legumes, feed conservation, natural pasture improvement (e.g., fencing, weeding), or the purchase of commercial feeds. Concurrently, changes in feeding practices may be associated with changes in herd composition (shift to smaller herds with crossbreeds), in herd management (gradual shift from free-grazing on own or communal land toward zero-grazing) and the associated additional labor inputs (feed preparation and conservation, ad libidum watering).

[4]http://www.parisschoolofeconomics.com/behaghel-luc/FTmaterials.zip.

[5]Heifer International is leading the program, with the participation of TechnoServe, ABS, ILRI (the International Livestock Research Institute), and ICRAF (the World Agroforestry Center). ICRAF's role in the program focuses on the dissemination of feeds and feeding practices.

We conduct the evaluation of the FT program during the expansion of the EADD program to new regions in Uganda, after it had been implemented in three countries (Kenya, Uganda, and Tanzania) during five years, serving about 40,000 farmers. Our research builds upon previous research on the first phase of EADD and the FT program, conducted by ICRAF and sociologists at Makerere university, and on the qualitative work we conducted together in 2012-2013. Based on in-depth interviews, focus groups and short surveys, these first investigations showed that while common feeding practices were widely disseminated, FTs had limited knowledge on the wider set of feed technologies. FTs may hence restrict themselves to teaching technologies that are directly relevant for them, raising the possibility that these do not correspond to needs of neighbouring farmers. This finding also raises the question on whether linking FTs with professional agents can enhance their knowledge base. Another finding was the importance of incentives such as social recognition for the FTs motivation. These findings, which echo findings in the broader literature, motivated the evaluation design with three variations that were refined during the pilot phase of the study.

In the first variation, selected professional extension agents (EAs) are contracted by the producer organizations (POs) to visit the FTs with two goals: providing technical support / additional training to FTs, and helping / monitoring them when setting up their training and visit work plan. They are instructed to review records kept by the FT. The EA themselves fill out reporting forms that are reviewed by PO executives. The visits by the EAs occur at the FT demand. More precisely, FTs selected by lottery are given twelve vouchers (allowing one visit per month during a year). They give one voucher to the extension agent to pay for the visit. The extension agent then redeems the vouchers at the PO. We will refer to this variation as the *linkage variation*.[6]

The second variation takes the form of a "thermometer display": a metal signpost is displayed at the FT's gate, with the picture of a healthy crossbred cow, the name of the FT and his or her phone number. In addition, space is provided on the board to show, for each month, the number of farmers who have benefited from training (number of trainees), the number of training sessions held, and the number of feed technologies implemented by the trainees and the FT. The information is filled by the FT. The number of feed technologies is obtained by counting the number of technologies from a check-list of 26 that has been distributed to the FT. We will refer to this variation as the *signpost variation*.

In the third variation, FTs receive an additional one-day training every 6 months, to train them to conduct needs assessment sessions with their trainees. They are provided with a method to follow. The final product is to write individual action plans with the trainees. We

[6] This variation in a sense tests the reverse mechanism to increase effectiveness compared to a study by Kondylis et al. (2017) in Mozambique. There contact farmers that had been initially trained by extension agents subsequently received direct central training, which led to no increased adoption among fellow farmers, even if the contact farmers themselves started implementing them. In contrast, we test whether regular visits by extension agents help reinforce the initial central training received from highly-skilled technical trainer-of-the-trainers.

refer to this variation as the *needs assessment* variation.

These three variations are meant to increase the effectiveness of the program. They were jointly designed by the program staff and the research team to answer operational and research questions, refined based on pilots conducted in 2013-14 in areas covered during the first phase of the FT program and on discussions with representatives of producer organizations and farmers in the Busoga region. They are at the same time clearly connected to economic theory and existing evidence on asymmetries of information and social motivation for service delivery, and on the role of modalities and heterogeneity in the diffusion of (agricultural) knowledge.

First, the variations aim to mitigate the consequences of asymmetric information and the related moral hazard issues in service delivery. Information is given to FTs, but FTs may lack the credibility or the motivation to transmit it to other farmers. By endorsing and publicizing their role, the signposts can increase credibility and provide a commitment device (directly inspired by Ashraf et al. (2014)). The linkage variation tests another approach, based on the monitoring by professional agents: their periodical visits and the associated reports provide incentives to the FT; they are also visible to fellow farmers, potentially increasing the FT's credibility.

Second, the variations test different input mixes in the production and diffusion of knowledge. The core FT program uses a mix of classroom and practical teaching, first given by ICRAF staff to FTs, then by FTs to fellow farmers. The needs assessment intensifies the former part, with one additional training session for FTs each semester. It also aims to promote a more active learning approach, by encouraging farmers to set individual goals, and to review them with the FT. The linkage variation diversifies teaching channels, by introducing the extension agent as an additional source of knowledge, likely complementing ICRAF teaching by relating it to animal health (most EAs have some veterinary training).

Third, the needs assessment variation is an attempt to take into account farmers' heterogeneity in social learning. It makes the heterogeneity salient by eliciting farmers' goals and constraints. It also raises FTs awareness against the risks of "one-size-fits-all" recommendations that may lead some to make unprofitable production choices.

It is important to note that, while being clearly related to hypothesized economic mechanisms, none of the variation provides a pure test of a given mechanism. Instead, one may expect each of them to act on several levers, not to mention unexpected effects that may arise. Importantly, the variations were designed in the continuity of the existing FT program, and are sometimes compromises between the hypotheses of program staff, farmers, and researchers. The linkage variation was in particular driven by the experience of program staff that backstopping by extension agents had been insufficient in previous phases. The needs assessment variation was primarily motivated by researchers' hypothesis on the importance of farmers' heterogeneity. Last, the signpost variation came as a tentative response (after piloting alternative responses in the field) to the intuition of all parties that some incentives were needed to sustain FTs' involvement, but that monetary incentives were at risk of backfiring by crowding out the intrinsic and pro-social motivations called upon by the program. Also, it was felt that direct monetary incentives would go against the nature of the FT program which is actually often named the "Volunteer Farmer Trainer program" by its promoters.

3 Evaluation design and data

The impact evaluation takes advantage of the expansion of the EADD program (including its FT component) to the Busoga region in Uganda in 2015-16.[7] While the other components of the EADD program developed on a business-as-usual basis, the FT component was randomly phased in across 77 geographical areas known as parishes.[8] This section details the different steps of the impact evaluation, and provides descriptive statistics on the sample and on treatment implementation.

3.1 Village recruitment and sample selection

The experimental sample is recruited from new parishes where the EADD program was expanding in 2014-15. A parish contains 8 villages, on average. Each village entering the program is associated with a potential FT. The recruitment of villages and potential FTs started in December 2014 and was completed by end of June 2015. It is a participatory process. Each recruitment "round" has three steps: (i) about 50-70 representatives (one per targeted village) are gathered by program staff and together establish criteria for selecting FTs; (ii) representatives go back to their villages, discuss with fellow farmers, and choose 2-3 candidates; (iii) these candidates come back to a new meeting with program staff, they are individually interviewed, and one candidate is retained for each village. We name him or her the "potential FT" as he will become FT once trained, which might occur in 2015 or in 2017 depending on whether the village belongs to the treatment group (entering the program early) or to the control group (entering late).

[7]The research was approved by the J-PAL Europe Institutional Review Board (first approval dated 2013-02-21, renewed annually). Moreover the research received clearance from the Uganda National Council for Science and Technology (research registration number A 519). Power computations were conducted in April 2015, once about a third of the rolling baseline survey had been completed. A pre-evaluation plan was finalized and registered on the American Economic Association Randomized Controlled Trials registry in July 2017 (before endline data collection), https://www.socialscienceregistry.org/trials/403.

[8]The research team had frequent interactions with the EADD leadership to ensurer that the roll-out of other components of EADD would not be based on the phase-in of the FT program. Monitoring in the field did not reveal any crowding-in or crowding-out of other EADD activities by the FT program. If anything, EADD activities took longer to develop than expected in all areas. Moreover, most of them were rolled out by large geographical clusters (counties or subcounties which are the appropriate level to organize producer organizations and milk bulking, for instance) that constitute strata in the lotteries assigning the FT program, so that most EADD activities are balanced by design across this study's control and treatment groups.

Once a village entered the evaluation, we ran a baseline survey among a sample of farmers and with all potential FTs. As the identification of villages and potential FTs was progressive, the survey was completed on a "rolling" basis between January and June 2015. In most of the 627 villages, five farmers were surveyed at baseline: the potential FT and four "other farmer." The later consist of three dairy farmers of his network, and one other dairy farmer who is not a member of his direct network. To establish the sample, a village leader was asked the list of dairy farmers in the village. On average, villages had about 30 dairy farmers. Then, the potential FT was surveyed first, and a specific module was included to identify his social network (with questions such as: name a farmer who has asked you for advice on dairy farming in the past three months, or to whom you talk at church). The three "network members" to interview and the "non-members" were then sampled using a list of random numbers. The goal of this stratification of the sample within villages was to distinguish two groups of dairy farmers: (i) farmers who are likely to become trainees of the FT based upon their pre-existing links with the potential FT; (ii) farmers who are less likely to become trainees (based on their absence of pre-existing link with the potential FT).[9]

3.2 Public lotteries

Two series of lotteries were conducted, first to select the treatment group, and second to allocate the different program variations.

The first lotteries selected about two thirds of the potential FTs (and the corresponding treatment villages) to start with the FT program early (year 1 of program expansion). The corresponding FTs are called "treatment FTs" (or simply, FTs, as they are the only FTs entering the program until the endline). Their villages are called "treatment villages" (vs. "control villages"). Lotteries were run in three waves (April 2015, August 2015 and September 2015) to limit delays between the progressive identification of potential FTs and the beginning of the program. In each wave, we ran several lotteries in different places (counties) yielding a total of 11 separate public lotteries. This allowed broad local participation in the process, but limited our ability to stratify the sample within each lottery. 6 of the 11 lotteries were stratified in two strata based on the characteristics of the FT (number of cows owned); 5 lotteries were not stratified further. In total, pooling lotteries together, we obtain 17 ($= 6 \times 2 + 5 \times 1$) strata. We account for stratification by introducing strata dummies in all our regressions.

These lotteries were conducted at the parish level, i.e. groups of about 8 villages are drawn together. This ensures that there is spatial variation in the implementation of the program: if every farmer had a FT nearby (in particular, if farmers from control villages all had a FT neighbor living in a nearby treatment village), then the evaluation would be threatened by inter-village spillovers – the distinction between treatment and control farmers

[9]For most of this paper, these two groups of farmers are pooled together, as we do not detect statistically significant differences. They are analyzed separately in Section 5.1.

would be blurred, with everybody more or less having access to the program in the first two years. Clustered randomization limits these spillovers and provides spatial variation to identify them.[10] Simulations conducted with baseline data showed that randomizing parishes strikes a balance between statistical power and the ability to identify spatial spillovers (see pre-analysis plan). Standard errors are clustered by parish throughout the analysis.

Table 1 describes the distribution of the sample across the treatment and control groups. The probability of assignment to treatment is close to 2/3.[11] By way of illustration, Figure 1 displays the map of control and treatment parishes in one of the two intervention districts. Local authorities (at the parish and sub-county level) were informed of the process, and the randomization was conducted in their presence (public lottery). Control villages started receiving the treatment just after the endline survey (fall 2017). Even though lottery losers expressed disappointment, the principle of the lottery was well accepted, as it provided a transparent way to manage over-subscription due to the limited ICRAF staff available to train the FTs.

A second round of public lotteries was then conducted to allocate variations. Each treatment FT personally drew whether (s)he would receive each of the three variations, with probability 1/2. Treatment FTs could hence be selected for up to three variations. Selection took place at the village/FT level; treatment FTs were invited to the meeting, informed briefly of the variations, and drew colored ballot papers from a box which determined their participation to each variation. These variation lotteries were run in two waves (August 2015 and January 2016) so that allocation to lotteries follows FTs incorporation in the program. Again, we ran separate lotteries in different places, for a total of 11 lotteries. In August 2015, these lotteries were held for FTs in the areas of four producer organizations, and in January 2016, lotteries took place in the areas of seven producer organizations.

Variations were thus cross-randomized across the 411 treated villages (or, equivalently, across the 411 FTs entering the program in year 1). As discussed in particular by Muralidharan et al. (2019), such factorial designs allow to test several treatments within a single experiment, but come at a cost in terms of interpretation. Our experiment was explicitly not powered to test for the presence of interaction effects between the three variations. Computations were made to check statistical power for pairwise comparisons, among the 411 treated villages, between those receiving vs. not receiving each of the three variations. Each of this comparison identifies the effect of a given variation, starting from a reference in which one fourth of villages had access to no other variations, one fourth to the two other variations, and the remaining half to either one of the two variations (Table 1).[12]

[10] Spillover across villages turn out to be of second order. They are analyzed in Section 5.1.

[11] Throughout the paper, "treatment" refers to the assignment to the FT program, irrespective of the assignment to the variations.

[12] Muralidharan et al. (2019) argue that such parameters may be of limited interest in randomized experiments where the policy relevant parameter is the direct effect of a given treatment, starting from a "business-as-

The addition of the variations makes the treatment (receiving the FT program) a composite treatment, with some villages receiving only the core program, and other receiving the core program plus one, two or three variations.[13] In practice, it is difficult to fully standardize a treatment in a field experiment where treatment is spread across 411 villages: one should expect a somewhat composite treatment anyway. The randomization of the variations can be seen as a way to control what components of the treatment are emphasized in each village. Statistical power limits the extent to which the impacts of each component and their interactions can be separately isolated; but the average effect of the FT program remains a well-defined and precisely estimated parameter.

3.3 Data collection and descriptive statistics

A Ugandan survey firm was contracted to collect the quantitative data, with a field coordinator linking with the research team throughout. The quantitative data was collected using computer assisted personal interviewing (CAPI). All farmers' houses were geo-located. The questionnaires used are shown in the Web appendix. The main sections of the baseline questionnaire are: households' composition and characteristics, livestock ownership and dairy production, land and agricultural activities, and social networks.

Table 2 displays descriptive statistics separately for farmer trainers and FTs.The sample consists of smallholder farmers, whose dairy farming activities are combined with staple and cash crops such as maize, bananas, cassava or sweet potatoes. On average, about 2/3 of agricultural land is devoted to crops, and 1/3 to pasture. Total land holding is on average 8 acres, with substantial variations (standard deviation is 15 acres among FTs, 13 acres among other farmers). A typical dairy farmer in the area holds two cows; a minority has one or two crossbred cows. To illustrate impacts along the causal chain, we build different standardized indexes following Anderson (2008) on milk production, technology usage, and technology knowledge.[14] Overall, FTs are more educated (about one third have completed high school), slightly wealthier (based on the number of assets held in a list of 24) than other farmers. They further differ from other farmers through their knowledge and usage of more cattle feeding technologies and the more frequent ownership of crossbred cattle, but they are not more experienced with cattle in general.

usual" situation in which no other treatment is received. In that case, they show, not surprisingly, that a more appropriate design "leaves the interaction cells empty." In the current experiment, we argue that identifying average effects of each variation in the presence of a mix of the other two variations is meaningful and useful, at least as a first step to un-bundle treatment effects. Indeed, none of the variation constitutes in itself a major deviation from the core FT program: variations are very much in the spirit of the overall program, but reinforce some of its aspects; moreover, the marginal costs of variations (as detailed below in the cost-benefit analysis) is relatively low compared to the cost of the core program.

[13] As noted above, we call the composite treatment "the treatment," or "the FT program." We call the treatment with no variation "the core treatment" or "the core FT program."

[14] The variance of the indexes is one, easing the interpretation of the differences between farmers and FTs.

Appendix Table A1 displays balancing tests using the baseline data.[15] Overall, baseline characteristics are balanced. For FTs, some coefficients appear statistically significant when estimated separately, but not when p-values are adjusted to account for multiple hypothesis testing following Romano & Wolf (2016). There remains a small imbalance among other farmers on milk production. To account for this, we control for baseline outcomes in our preferred specifications (see Section 4).[16]

Appendix A.1 provides additional details on midline and endline data collection. Importantly, due to intensive tracking, attrition at endline was very limited. The overall response rate is 98.2%; it is 98.3% among control farmers (28 non respondents), 99.2% among treatment farmers (7 non respondents); 96.8% among control FTs (13 non respondents), and also 96.8% among treated FTs (7 non respondents).

3.4 Program implementation

Monitoring data indicate that compliance was high for the treatment, and while some FTs stopped training after a few months, most held multiple trainings with the intensity of training activity more or less constant over time. The survey data further confirms that the program had a sizable impact on trainings received and given by FTs. Compliance was also high for the variations (in terms of participation), but not all FTs reported making active use of tools provided in the signpost or needs assessment variations (i.e. a substantial share did not update signposts with training activity or used the needs assessment forms). See Appendix A.3 for details.

4 Main results

Given random assignment to treatment within strata, we estimate simple OLS regressions of the form:

$$y_{ijs}^{gt} = \alpha^{gt} + \beta^{gt} T_{js} + \delta_{gt}^{'} Y_{ijs}^{g0} + \gamma_{gt}^{'} Y_{ijs}^{g0}.T_{js} + \mu_s^{gt} + \varepsilon_{ijs}^{gt}. \tag{1}$$

In this equation, y_{ijs} is an outcome variable (e.g. milk production) measured for farmer i in village j which belongs to lottery stratum s. The regression is run primarily on endline data (indexed by $t = 2$), but some measures of take-up and intermediate outcomes are analyzed with midline and endline data ($t = 1, 2$, respectively). Potential FTs ($g = 1$) and other farmers ($g = 2$) are analyzed separately. T_{js} is the binary treatment variable of interest: it is equal to 1 if village j is in a treated parish, 0 otherwise. μ_s^{gt} is a stratum (fixed) effect. ε_{ijs}^{gt} is the error term. Given the clustered randomization design, standard errors are clustered by parish.

[15] The table presents estimates of equation (1) without covariates.

[16] The imbalance is driven by the early lotteries conducted in April 2015, where baseline data had only been recently collected and was imperfectly used for stratification. The imbalance had been noted by the time the pre-analysis plan was written, and we pre-committed to a preferred specification controlling for baseline values of the outcomes, whenever available.

Whenever lagged values of the outcome are available, we include them as control, following equation (5.4) in Athey & Imbens (2017). Specifically, we discretize the lagged outcome variable into five indicator variables corresponding to the five quintiles. We demean these variables, interact them with the treatment indicator, and include them as controls (denoted Y_{ijs}^{g0} and $Y_{ijs}^{g0}.T_{js}$) in equation (1).

β^{gt} gives the average impact of the FT program at time horizon t for subsample g. It is an "intention-to-treat" (ITT) parameter, as actual exposure to treatment varies across FTs depending on attendance to the training refreshers, and across farmers depending on how active the FT is and on the demand for training by the farmer. The impact of the FT program on non-FT farmers (β^{2t}) is a mix of the impact of actually becoming trainee and of not becoming trainee but possibly benefiting from spillover effects within the village (e.g., receive advice from your neighbor who is a trainee). As mentioned in Section 3, attrition is extremely low throughout the experiment, so that we do not implement any correction. Spillovers between villages are analyzed in details in Section 5.1; we do not find evidence that they are of first order. Presenting most results in a model without spillover across villages allows us to get less noisy estimates; if anything, the resulting estimates can be viewed as lower bounds of the complete effect.

Table 3 shows the program impacts on four outcome indices[17] along the hypothesized causal chain: feed and feed practices knowledge, feed and feed practices adoption, milk production, and profits.[18] We find large positive impacts of the program on farmer trainers and other farmers. Standardized effects on knowledge and adoption are particularly large for FTs, and highly significant. As expected, they are much smaller for other farmers: transmission of knowledge is imperfect. Yet a 0.08 standardized effect is sizable if compared, for example, to standardized effects found in the evaluation of educational interventions on test scores; this is all the more striking given the small treatment doses received by the average farmer. These effects are statistically significant at conventional levels; they remain so for FTs after adjusting for multiple hypothesis testing over the family of four indexes following Romano & Wolf (2016); for farmers, the effect on knowledge is no longer significant (adjusted p-value = .22), but the effect on adoption is (adjusted p-value = 0.06). Effects on milk production by other farmers are economically meaningful and statistically significant even after adjusting the p-value to control the family-wise error rate. Impacts on milk production of FTs, and profits of both FTs and other farmers are positive, but less precisely estimated. Milk production and especially profits are indeed hard to measure, which may explain the large standard errors.

[17]Following Anderson (2008), we construct four corresponding indexes as weighted sums of individual variables, with weights inversely proportional to the variance-covariance matrix of those variables (see Appendix A.1). We normalize each index dividing it by its standard deviation in the control group, so that coefficients (β) can be read as standardized effects.

[18]"Profits" refer to net revenues from dairy activities; more details on breakdown and measurement are given in Table 6 and Appendix A.1.

As noted in Section 2, program effects estimated in model (1) are composite effects of the FT program with or without the linkage, signpost, and needs assessment variations. The evaluation was not powered to separately evaluate the impact of the eight combinations (with 0, 1, 2 or 3 variations). Instead, we estimate the following OLS regressions in the subsample of treated villages:

$$y_{ijs}^{gt} = \alpha^{vgt} + \beta^{vgt} V_{js}^{v} + \delta_{vgt}^{'} Y_{ijs}^{g0} + \gamma_{vgt}^{'} Y_{ijs}^{g0}.V_{js}^{v} + \mu_{s}^{vgt} + \varepsilon_{ijs}^{vgt}, \tag{2}$$

where V^v is an indicator variable that is equal to 1 if the village (FT) was assigned to variation v (where v denote linkage, signpost, or needs assessment). The equation is estimated separately for each variation, as justified by orthogonal assignment to each of the variations. Note that standard errors here do not need to be clustered by parish, as randomization of treatment variation is at the village level.

Table 4 displays the results of model (2) on the same outcomes as Table 3. Effects are less precisely estimated, and point estimates are smaller than overall program effects, with notable exceptions. We first focus on the linkage and signpost variations and defer the discussion on the needs assessment results to the next section. Both the linkage and signpost variations have large and statistically significant impacts on the milk production and the profits of other farmers, but they have no impacts on their knowledge and limited impact on technology adoption. Hence the linkage and signpost variations appear to impact other farmers in their productive decisions and effectiveness, even though their knowledge did not improve: this may indicate that their trust in FTs' advice is enhanced, so that they react more strongly to information. We first discuss robustness of these main findings, and then return to evidence on such mechanism below.

4.1 Robustness

Unsurprisingly, results are robust to minor specification changes such as removing controls for baseline outcome values (see Appendix Table A2). More importantly, we check that the results based on self-reported measures are corroborated by objective measures of impact. Indeed, given the nature of the intervention and the public lotteries, self-reported outcomes may suffer from systematic response bias based on treatment status (social desirability bias). We therefore implemented a separate complementary data collection at endline to record milk quantities at milking time and test milk quality, as well as observed animal health (all key indicators of improved feeding practices).[19] Observable measures of these key outcomes were

[19] More specifically, 10 para-veterinary technicians were hired, who attempted to visit one (potential) farmer trainer and one other farmer per day, during the morning and the evening milk sessions for 60 days. The visits took place without advance notice and were programmed to occur a few days after the enumerator visit for endline data collection. The para-veterinary workers measured milk quantities, collected a small quantity of milk for purposes of testing the quality with a lactoscan (the testing was subsequently done in the base camp), measured the animals' weight (using a weight band), and noted other observable indicators of animal health.

obtained for 420 FT and 518 farmers. Only a subsample of FTs and farmers were sampled for these visits, given the need to have the timing of visits exactly around milking times (in particular during very early morning hours). The final sample of this analysis is balanced. For all visited farmers, separate measures were obtained for indigenous and cross breeds (where farmers had both). Health and weight measures were obtained for all milking cows, as well as for two randomly selected non-milking cows by breed. However, as the breed composition is endogenous, and given evidence of program impacts on breed composition (see further), we pool results of health and size measures for cross breeds and indigenous cows together. We report results with and without controls for the para-veterinary technicians that measured the outcomes.

Results with the observed milk quantities corroborate the self-reported ones (correlation coefficient equal to 0.7, and estimated program impacts are qualitatively similar though less precise as observed quantities are only available for a subsample). Table 5 shows a significant increase in average cow weight, with the impacts being more important (and significant) for the farmers (compared to the farmer trainer). The increased weight corresponds to about 15-20% of a standard deviation, reflecting the findings for milk production.[20] On the other hand, we find no consistent significant results for indicators of bad health – and if anything some evidence of worse health outcomes. The latter are driven by symptoms that are likely unrelated to feeding practices (abnormal discharges, swollen lymph nodes, etc.). They possibly reflect changes in herd composition given the higher susceptibility of cross-breeds to health hazards. Overall, this suggests that impacts on productivity are not driven by improvements in animal health. Finally, Table 5 finds no impact on milk fat content, another potential indicator of improved feeding practices, and in particular of more balanced cow diets.[21] The results of the differential impacts of the variations in Table A3 does suggest, however, that the signpost variation led to an increase in fat content for the farmer trainer. The impact is relatively large (.26 sd) and robust to various specifications and robustness checks.

Turning to robustness for profits, Table 6 breaks down the measure of profits used in previous tables. The construction of profits from dairy farming involves several difficulties. It is not fully clear whether feed expenditures are well reported and how to handle (net) investments through the acquisition and the sale of livestock that should be amortized over several years.[22] As shown by the standard errors, these two elements add noise. The most

[20]Appendix Table A3 shows that impacts are larger when the FT is assigned to the linkage and signpost variations. Overall, these results hence confirm the productivity results.

[21]Estimates are shown for the raw fat content, for standardized fat content after taking out breed and session fixed effects (as fat contents of cross breeds differ from those of indigenous cows, and morning and evening milk also systematically differs), and for estimations that exclude milk samples for which there was concerns about quality (related to temperature of testing, duration between collection of sample and test, or non-standardized treatments of the samples, etc.).

[22]Piloting of the endline questionnaire showed that farmers had difficulties recalling their purchases of feeds. As such purchases are irregular and mostly took place during the dry season, we asked about purchases in the past three months, which we multiply by two. Surveys took place during one of the two yearly dry seasons.

conservative specification includes them without amortizing livestock sales and acquisitions (bottom line). As robustness checks, we display the results excluding these two elements, or excluding livestock sales and acquisitions.[23] All alternative computations suggest positive impact on net revenues from dairy farming, even though the estimates are not precise, and the point estimates depend on how we treat livestock acquisition and sales, as well as feed purchases. This confirms the difficulty of computing annual profits based on a simple survey, and justifies our focus on robust measures of milk production. Note that, in monetary terms, impacts on profits remain modest. Even the favorable estimate of 152,000 Ugandan shillings per year (excluding livestock investment) translates only into about 3 US dollars per month.[24]

Appendix Tables A4-A6 further show robust evidence that the linkage and signpost variations increase the farmers' profits. This results from the large increases in milk production observed in Table 4 that are not offset by the small increases in expenditures.

4.2 Mechanisms

The basic causal chain — i.e. the program increases the knowledge of FTs, most FTs train actively and pass on knowledge to their potential trainees, spurring technology adoption and gains in production and profits – is validated by the evidence above.

We now examine the evidence on the detailed mechanisms driving these effects. For this, we consider the effects of both the main treatment and the variations on intermediate variables. We focus on information collected at the endline, but also mobilize data from the midline when relevant and account for variations in take-up of the program. We distinguish four sets of mechanisms related to (a) access to technical information, (b) the incentives of FTs, (c) social interactions between farmers, FTs and EAs, and (d) complementarities with other components of the EADD program.

Access to technical information

A major constraint to volunteer farmer extension schemes could be the quality of the technical information being diffused. To address this constraint, the linkage variation sought to provide FTs with more regular access to technical information. Using information on ways farmers learnt about different feeds and feeding practices, Table 7 shows that all treatment FTs (not just those benefiting from the linkage variation) received more frequent advice from professional extension agents and veterinaries. In villages with the linkage variation, FTs learnt on a

Many farmers, however, do not purchase any feeds.

[23]Note that we do not take into account changes in household labor inputs, which are not obvious to value and on which we find little impact (results not reported here).

[24]Accordingly, we do not find consistent impacts on welfare outcomes (food diversity, enrollment of kids at school, health of household members). This is consistent with the relatively small size of the effects on profits and the fact that such outcomes take time to change.

larger number of feeds from an agent. Access to technical information, an important condition for the intervention effectiveness, was hence reinforced by the linkage variation.

Incentives and motivation of FTs

The incentives of agents are a central concern in the analysis of extension interventions. In the volunteer farmer training model, FTs are expected to be driven by pro-social motivation, but could as well respond to some forms of accountability to their communities, and to monitoring by extension agents. They may also be motivated by the learning opportunities offered by the program or by gains from social capital or recognition. The reports of both potential FTs and other dairy farmers suggest that the program significantly increased the pro-social motivation of potential FTs, and this largely fueled their training activity and its efficacy. We directly measure pro-social motives, among other sources of motivation, by asking FTs to weight the reasons that make them train others.[25] As shown in Table 8, the program increases the weight potential FTs give to the "importance of helping others as driving effort" (compared to gaining knowledge through training, being known, receiving inputs, deriving other benefits). This is confirmed when asking farmers about the motivation of the FT of their village. We also asked FTs about satisfaction derived from pro-social behaviors (asking about sources of satisfaction though questions such as: "In the last 7 days, how many days have you been happy because you could provide somebody with advice?" or "In the last 7 days, how many days did you feel good because you were able to help other farmers?", and taking the first factor of a principal component analysis of their correlations). Treatment FTs exhibit higher levels of pro-social motivation.[26]

FTs were selected for training through a participatory process and, while they were not committed to any targets, the accountability to their communities can also fuel their activity. We do not have a direct measure of accountability and we cannot distinguish it from extrinsic pro-social motivation (associated with public recognition). However the effects of the signpost variation can reflect such incentives (even if only about half of farmer trainers updated their training activities on the signposts). We do not observe that signposts increased training activity significantly, even though they might have increased the actual effectiveness of their training and farmers' perceptions of FTs' advice usefulness, as implied by the variation's impact.

The linkage variation involves more formal monitoring as professional extension agents not only provided FTs with more regular access to technical information but also observed their training activity. Our results confirm that extension agents more likely participated to the training sessions of FTs in the linkage variation (a 6 percentage point increase). Such extension agent involvement could help explain that FTs in the linkage variation more often

[25] The enumerators used a beans allocation scheme to materialize the weights.

[26] Table 8 shows limited evidence that variations increase pro-social motivation, yet the point estimates on FT's pro-social motivation (as reported by farmers) are consistently positive and large.

remained active for a longer period.

FTs might also respond to some more direct benefits from the program. A share of them mention access to knowledge as a source of motivation and this could explain some of the effects of the linkage variation.[27] Overall, we find evidence in support of different sources of motivation: pro-social motivation, extrinsic motivation of public recognition or accountability checks, monitoring by extension agents and access to knowledge. While the signpost variation can have enhanced public recognition or accountability checks, the linkage variation reinforced monitoring and knowledge access.

Social interactions in dairy farming and credibility of FTs

The program rests on farmer-to-farmer diffusion of technologies through formal training but also though broader and less formal social interactions We do observe changes in the social interactions and relationships between potential FTs and their trainees. Table 9 shows that, in treatment villages, both potential FTs and farmers were considerably more likely to have talked about dairy farming to another villager, or asked him advice about dairy farming (a 26 and 7 percentage point increase for FTs and farmers compared to 51% and 22% in control villages). Treated FTs also talked to a larger number of fellow farmers of their village about dairy farming than control ones (to 5.3 people on average compared to 2.5 in control villages - the corresponding effect for farmers is not statistically significant). These interactions seem to be reinforced by the linkage variation.

The credibility of FTs, and potential lack of confidence of fellow farmers in them, constitutes an important potential constraint to these social interactions. The linkage and signpost variations could address this constraint by publicly recognizing FTs' activity. Table 9 shows that the program increased farmers' perceptions of how knowledgeable the FT of their village is regarding both feeding practices and other aspects of dairy farming. The signpost variation seems to have increased a little more farmers' perceptions of FT's knowledge (in other aspects of dairy farming than feeding), but the linkage variation had no significant effects here (access to more knowledge from extension agent seems more important).

Overall, the effects of the main program appear to go through more intense social interactions related to dairy farming, with FTs but also broadly between farmers, and the linkage variation contributes to these interactions, by introducing more knowledge and broadly supporting the training and the credibility of the farmer trainers.

[27] While they might also value the recognition they obtain from their activity, we find no evidence that the program affected FTs or farmers participation in local groups (note that 90% of potential FTs and about 40% of dairy farmers already participate in those groups) or in community-based credit or savings groups (respectively about two-thirds of potential FTs and 50% of farmers participate in those groups). See online appendix A8.

Complementarities between FT program and other interventions

Another set of mechanisms relates to the complementarities between the FT program and several other interventions of the EADD program, including the distribution of inputs (mainly seeds), the promotion of crossbred cows, and access to other sources of support (extension agents, other community-based facilitators of dairy interest groups and agro-dealers, all of which were introduced in both the treatment and control parishes of our study).Where the program operates, as seen in Table 10, FTs do receive small quantities of seeds and in turn distribute some seeds to their fellow-farmers. In addition FTs seem to have received more seeds in villages with the linkage variation, suggesting that the extension agents intervened here.

Another recommendation passed through the FT program and reflecting the broader messages of EADD is the shift toward smaller herd with crossbred cows. Table 10 provides some evidence that such messages may have been effectively put into practice, facilitated by the FT program, with a shift from indigenous to crossbred cows among farmers. The increase in the number of crossbred cows held by farmers is statistically significant and meaningful (more than a 40% increase). The shifts are small and not significant for FTs, possibly because they were already more likely to have crossbred cows even in the control.

Hence, while the promotion of crossbred animals was an important aspect of the entire EADD program and not only its extension arm, the FT intervention seems to have spurred complementarities between changes in herd compositionand adoption of new feeding practices.

4.3 Additional qualitative evidence

Feedback from structured interviews and focus group discussions further help understand the causal pathway between changes in feeding practices and milk productivity. It also confirms that increases in milk productivity were perceived to be important by the farmers and farmer trainers, as they specifically pointed them out, and indicated several factors leading to such changes. Farmers and farmer trainers reported, for instance, that better animal feeding practices (e.g. adding salt to water for animals, timely watering of animals, use of feeding troughs, feeding animals on crop residues and hay during dry seasons) made the biggest difference, leading to increased milk yield. Respondents reported a positive appreciation of the specific farmer-to-farmer method underlying the FT approach (proximity of the FT; increased confidence when the FT received visits by extension agents). The interviews also revealed, however, that several practices were not applied for various reasons: they were too expensive (multi-nutrient mineral blocks, silage method of pasture conservation, water harvesting) or technically difficult to implement (silage), or required inputs that were not accessible (some seeds or planting materials).

Qualitative insights were also crucial to disentangle some of the mechanisms underlying the differences in impacts for the different variations. They indicate for instance that the signposts served as a reminder not just for the farmer trainers, but also for the farmers, and further increased the credibility of the farmer trainer.[28] Similarly, the qualitative work confirmed that the linkage variation was implemented as envisioned. Extension agents came on appointment and were given a voucher only if they provided training and support of at least one hour. Activities engaged in with extension agent included guidance on filling in of report formats, consultation, teaching/coaching, and training other farmers. The extension agents were instrumental in reinforcing the messages of the FTs. Many FTs organized the extension agent visit to coincide with their group training meetings so that the extension agent co-facilitated the FT training. This added clout to the FT making fellow farmers to believe him/her as a trainer: "I am now seen as an important person because every month people see the motorcycle of the extension agent parked at my home!"

5 Heterogeneity and knowledge diffusion

As discussed above, the needs assessment variation was designed to enhance the capacity of the FT program to produce customized recommendations to heterogeneous smallholders. Importantly, the core FT program already tried to adapt information and training: its first step when entering the Busoga region was to conduct a global assessment of dairy farmers' training needs, which resulted in identifying a set of about 30 feeds or feeding practices to which FTs were progressively trained, and which were presented as a menu to fit the different needs of the farmers. In addition, by training one FT per village, the program ensures that the person training the farmers is aware of local needs. The needs assessment variation pushes that logic one step further, by training FTs to conduct needs assessment at the individual farmer level. As shown in Table 4, however, this variation has no detectable impact on farmers' knowledge, adoption or production, even if it significantly increases FTs' knowledge. By adding two additional sessions per year and broadening the set of technologies covered, the needs assessment variation intensifies the FT training. But the additional knowledge is not transmitted to fellow farmers; and it does not impact FTs' own practices or production in a detectable way. The customization variation hence did not work as envisioned. The qualitative work confirmed that needs assessment prior to training by the FTs generally did not occur, even if needs assessment training of FTs occurred, and FTs appreciated this additional training. This is consistent with the limited reported use of the needs assessment forms in the quantitative survey. This may be due to the difficulties faced by FTs to differentiate training contents by farmers, or to the difficulties of farmers to fill individual work plans, rather than to the absence of heterogeneous needs of farmers.

[28]For instance, farmers reported that the photos of cattle on the signpost motivated them to engage in better animal feeding practices so that their cattle can look good and healthy like the ones on the signpost. Farmer trainers with signposts were more trusted than those without them – a finding that suggests that the positive marginal effect of the signposts may be partly due to a negative effect on FTs without signposts.

The fact that a variation aimed to take individual farmers' heterogeneity into account did not work as intended is not sufficient to conclude that smallholders' heterogeneity does not matter in this context. In particular, it is possible that the core FT program is sufficient to take care of the existing heterogeneity. Indeed, the involvement of the community in the selection of FTs may ensure that each FT is sufficiently representative of his village, or otherwise capable of catering to local needs. This may be reinforced by assortative matching, if the FT mostly trains (and is called upon by) farmers facing similar farming conditions. This section provides empirical evidence to test such an interpretation. Specifically, we ask three questions. Fist, how much heterogeneity is there initially across and within villages, and does this condition information diffusion about cattle feeding? Second, is there assortative matching between FTs and their trainees, that acts as a channel for customization? Last, does the program have heterogeneous treatment effects?

5.1 Baseline heterogeneity and program spillovers

The baseline descriptive statistics in Table 2 indicate that dairy farmers in the Busoga region differ in several important dimensions: some have crossbred animals, others do not, some have large herds and pasture land, some just have a few animals and no pasture, a few use advanced feeding or breeding technologies, etc. From the perspective of the FT program which trains one FT per village, an important question is whether this heterogeneity is mostly between villages (in which case the choice of a local FT may be instrumental in adjusting training to farmers in his village) or within villages (implying that further customization has to be ensured by the FT for diverse village members). Table 11 provides this decomposition on the main variables of Table 2. Specifically, we report the intra-cluster (or intra-village) coefficient of correlation for four baseline indexes (household education, knowledge of feeds and feeding practices, adoption of these feeds and feeding practices, and milk production) and for the number of crossbred and of indigenous cows. The heterogeneity across villages is low for education, milk production, and number of cows (points estimates below 0.14, upper bound of 95% confidence interval below 0.18). But it is larger for the knowledge and the diffusion of feeds and feeding practices. Note that the share of variation between villages is even larger if we exclude FTs from the sample (as already noted from Table 2 and further detailed below, FTs differ significantly from other farmers, driving some within-village heterogeneity). Table 12 similarly shows intra-village correlations for key farmer characteristics, that are hypothesized to be potential important drivers of heterogeneity in adoption. None of these are very high (point estimates ranging from 0.07 to 0.29), indicating that even within a village, farmers differ substantially in terms of their productive assets, experience or labor allocation to dairy.

Nevertheless, the heterogeneity across villages in terms of technology diffusion may imply that information and training needs differ too, justifying the training of one FT per village.[29]

[29]The treatment did not, however, increase the within-village correlations of the main indices (see Table

It may also limit the room for knowledge spillovers across villages, in particular of neighbouring villages differ in their training needs. Table A12, which shows that the intra-parish correlation coefficients for knowledge and adoption of the diffusion of feeds and feeding practices are much lower than those at the village level, suggest this could be the case. We know turn to a more in depth analysis of such spillovers.

Knowledge diffusion and spillovers

Spillovers are the engine of the contact farmer model: by training a few, the goal is to transmit information to many. The FT program trains one farmer per village. Villages in Busoga are well-defined administrative entities, and constitute pools of about 30 dairy farmers on average that can be served by a farmer trainer. Yet, many forms of social and dairy-related interactions – such as markets or religious meetings – occur at higher geographical levels, e.g. parishes. It is therefore important to analyze whether the program has spillover effects across villages.

The clustered randomization design creates random spatial variation in the exposure of farmers to treated villages in their proximity that allows to identify spillovers between villages. Following our pre-analysis plan, we use two measures of exposure to spillovers. The first measure is the number of FTs who live within a 2 km distance from a given control farmer, denoted $NT02$, and the additional number of FTs who live between 2 km and 5 km away ($NT25$). These numbers vary greatly, depending on whether the control farmer we consider lives close to the border of his parish, and whether the neighboring parishes are treated or not. For instance, control farmers live less than 2 km away from 0 to 8 active (treated) FTs (with an average of 1.1); and between 2 to 5 km away from an additional 0 to 36 (with an average of 10.9). We estimate the following model:

$$\begin{aligned} y_{ijs} =& \beta_0 + \beta_1 T_{js} + \gamma_{02} NT02_{ijs} + \gamma_{25} NT25_{ijs} + +\zeta_{02} NT02_{ijs} \times T_{js} \\ &+ \zeta_{25} NT25_{ijs} \times T_{js} + \delta_{02} N02_{ijs} + \delta_{25} N25_{ijs} + \mu_s + \varepsilon_{ijs} \end{aligned} \tag{3}$$

where $N02$ (resp. $N25$) are the number of potential FTs within a 0-2 km (resp. 2-5 km) band around the FT – to account for the impact of population density, in the same spirit as Miguel & Kremer (2004).[30]

A second measure of exposure indicates whether there is at least one active FT in the neighborhood (2 km radius) of a control farmer. The idea is that the intensity of exposure (captured by the number of neighboring FTs) matters less than having access to at least one FT nearby. We denote by $T02$ the corresponding indicator variable. Again, this measure varies substantially: about half of control farmers (50.5%) live less than 2 km away from an

A13).

[30] $N02$ and $N25$ capture unobserved differences across villages located in more or less dense areas, which are, by design and in expectations, the same for treated and control villages. Hence no interaction term with T is needed. The model with the additional interaction terms yield qualitatively similar results.

active FT. The estimated model writes:

$$y_{ijs} = \beta_0 + \beta_1 T_{js} + \beta_{02}(1 - T_{js}) \times T02_{ijs} + \mu_s + \varepsilon_{ijs}, \tag{4}$$

where $T02$ is an indicator variable equal to 1 if there is at least one treated FT withing a 0-2 km band around the FT. This is a model of radical contamination: basically, it assumes that it is enough to have one treated FT within 2 km in order to receive full spillovers.

Table 13 shows the main coefficients from equations (3) and (4), labelled as model 2 and model 3, respectively. Model 1 recalls the results without spillovers (it corresponds to equation (1) without covariates). The table shows very limited evidence of spillover effects across villages: the coefficients on $NT02$, and its interaction with treatment status T are not statistically significant, while the coefficient of $NT25$ is significant for adoption and milk production, but not for knowledge and profits, and neither is any of the interaction effects. Moreover, to the extent there is any spillover, it is limited in size. To quantify average spillover effects, one needs to combine the estimated coefficients with the intensity of exposure to spillover. This is denoted as the "implied spillover effect": for instance, according to model 3, the average spillover effect is to reduce knowledge scores by 0.3% of a standard deviation, a number computed as the weighted average of the effect on treated farmers (2/3) and control farmers (1/3) for mean exposure levels. This implied spillover effect is not statistically different from 0 (p-value = 0.877), and small compared to the "implied total effect" summing the spillover effect and the direct effect of the program on treated farmers (5.6% of a standard deviation). The overall picture is that we do no detect statistically significant spillover effects across villages (p-values in last line of the table), and that these effects seem limited, with implied spillover estimates accounting at best for a limited fraction of total effects.

The absence of (detectable) spillover effects across villages may come as a surprise, given the quite strong diffusion of knowledge and adoption within villages, and the fact – documented in Section 5 – that farmers who are not reported as members of the FT's network do not show different treatment effects compared to network members. Does diffusion stop at village boundaries? To dig deeper, Appendix Table A14 analyzes the diffusion of knowledge, the attendance to training led by the FT of a neighboring village, and the diffusion of seeds. At the endline survey, farmers were asked to define what "Katiti" was. Though it sounds familiar, the word Katiti does not exist in the local lusoga language, and was purposefully invented to baptize one fodder shrub (Gliricidia Sepium) that the FT program trainers promoted during the training given to FTs. It acts as a "marker" to signal the diffusion of information taught during these trainings: a farmer correctly answering that question must have been, in one way or another, exposed to some information from the program (unless he makes a good guess). Somewhat surprisingly, results in model 1 show that 9.7% of farmers in control villages can define Katiti. But the direct program effect is large: an additional 4.7 percentage points in treated villages know the correct answer (almost a 50% increase). While the number of control farmers responding correctly suggests information diffusion across villages, the results of models 2 and 3 do not support that hypothesis: we do not find significant spillover effects in these models. We next look at whether a farmer has received training in dairy farming over the past 12 months. This proportion doubles for farmers in treated villages. Yet, the fact of having more treated FTs living in nearby villages does not significantly increase that probability further (models 2 and 3). Similarly, the probability of receiving seeds and seedlings from a fellow farmer to plant cattle feeds is much stronger among treated farmers than among control ones, suggesting that the FT distributes the seeds, but that they do not diffuse beyond the village (or have not yet diffused).[31]

Table A19 provides more detailed evidence on the propensity of farmers to look for training outside of their village. The endline survey had asked farmers whether they attended training not only with the (potential) FT in their village, but also with the 24 closest (potential) FTs. We look at training incidence and number of training sessions for each of the corresponding farmer - potential FT pair, controlling for distance between the farmer and the FT. Very strikingly, training is much more frequent when the FT and the potential farmers live in the same village (increasing training incidence by 18.5 percentage points for a control FT, and by $18.5 + 26.0 = 44.5$ percentage points for a treatment FT). By contrast, a treated FT does not have significantly higher chances to train a farmer of a neighboring village than a control (potential) FT. In short, potential and actual FTs mostly train farmers of their own village. This remains true for non-network farmers. It also remains true if we consider FTs with signposts, although they could be more visible to farmers from neighboring villages.

To summarize, spillover effects across villages do not seem to be a first order phenomenon here compared to the strong diffusion of information and practices within villages. The evidence suggests that by putting the FT in charge of a given village, the program has led to information diffusion mostly within villages. These patterns are also consistent with FTs being well placed to target relevant messages to people within their village (about whom they may have more local knowledge) but less able to reach people outside local circles. Either way, they arguably validate the FT approach of training one FT in each village.

5.2 Assortative matching?

Given the evidence that FTs provide training only to village members, the next question is whether they train all of them, or whether there is "assortative matching" through which the FT only trains the subset of dairy farmers who share similar farming conditions.

To investigate such assortative matching, we analyze the characteristics of farmers, in a given village, along two dimensions: first, we compare members of the FT network with non-

[31] Appendix Tables A15 to A18 replicate the analysis of Tables 13 and A14 splitting the sample of farmers between those who were sampled in the network of 10 farmers reported by the FT at baseline and those who were randomly chosen from a list of dairy farmers in the village, excluding the FT network farmers. It may indeed be the case that farmers only get connected to neighboring FTs if they do not know an FT (or potential FT) in their village well enough. However, we do not find large differences.

members of that network; second, we compare trainees of the FTs with non trainees. The first distinction is made by the FT himself in the baseline survey, when he is asked to designate 10 dairy farmers with whom he would interact most. As explained above, we then sampled three network members and one non-member per village. The second distinction is reported ex post by the farmers during the endline survey, where we ask them whether they have received training from the FT over the previous 12 months. Table 14 compares the baseline characteristics of these groups of farmers by running linear probability models.[32] The separate regressions displayed in the different columns thus show which characteristics predict which farmers are members of the FT network (panel A), and which farmers will become a trainee of the FT (panel B). We consider as potential predictors the whole set of variables displayed in Table 2. For clarity, we only list the coefficients on the three indices concerning knowledge, adoption and milk production; the predictive power of the other variables is summarized by the increase in adjusted R^2 when they are included. For instance, column (1) in Panel A shows that farmers who produce milk more are more likely to be members of the FT network; knowledge and adoption are not statistically significant predictors. Importantly, the predictive effect of milk production is small: a one-standard deviation increase in milk production increases the probability to be a FT network member by 3.5 percentage points, from a 75% baseline. Accordingly, the adjusted R^2 of the model is low (0.005). Adding all the other variables listed in Table 2 barely increases the model's predictive power. To sum up, members of the FT network do not differ much, in terms of observable characteristics, from non members. Columns (1) and (2) in Panel B yield a similar picture for trainee selection: trainees of the FTs do not differ much from non trainees.[33] Panel C in Table 11 provides a benchmark to assess the predictive models of Panels A and B: we check whether the same characteristics predict who is selected as an FT (vs. other farmers). Most predictors enter significantly, and with large coefficients; as a result, the adjusted R^2 reach values 10 times larger than in Panels A and B (0.103 with the three indexes, and 0.216 with all the covariates listed in Table 2). To summarize, FTs systematically differ from other dairy farmers; yet they do not seem to match with specific farmers in their villages.

It could still be the case that FT network members or FT trainees look on average the same as other farmers on the whole sample of 627 villages, but that, within a village, they look more similar to the FT. To check this hypothesis directly, we now specify all the predictors in terms of distance to the FT, that is, we subtract from the characteristics X_{ij} of farmer i in village j the value of that characteristics for the FT in village j, $X_{\mathrm{FT},j}$. We then use the distance $|X_{ij} - X_{\mathrm{FT},j}|$ as predictor. Results in columns (3) and (4) of Panels A and B provide

[32] Just as for balancing tests, a regression of the selection (treatment) indicator on the baseline variables provides a convenient summary of the (partial) correlations.

[33] Membership in the FT network and participation to his training sessions are correlated (network members are 20 percentage points more likely to report receiving training from the FT). But the correlation is imperfect; the characteristics predicting these two statuses in Panels A in B can therefore differ.

no evidence that members of the FT network look more similar to the FT than non members, or that FTs are more likely to train farmers that are similar to them.

Overall, we find no evidence that, within a village, FTs are more connected to, or tend to attract as trainees, farmers that face similar farming conditions to them. This may not be so surprising. Network membership likely follows other lines, including family links. And the fact that a large and representative share of dairy farmers in the village receives training is consistent with the FT program instructing FTs to train all farmers in their village. Qualitative interviews and focus groups confirm that FTs felt they had to train all farmers in their village, and that all farmers felt entitled to their support.

5.3 Treatment effect heterogeneity

To further analyze the possible importance of heterogeneity, we test whether either a theory-driven or a data-driven approach reveals heterogeneous treatment effects. If we were to find heterogeneous treatment effects, that could be an indication that the program is insufficiently customized to take into account the heterogeneous needs of the farmers. Conversely, homogeneous treatment effects could reveal perfect customization, or no need for customization.

The rich baseline data at our disposal allows to test for heterogeneous effects in a variety of dimensions, but this comes at a risk of data mining and false discoveries. To discipline investigations, we use two complementary approaches: first, we interact the treatment variable with a handful of variables that had been pre-specified in a pre-analysis plan based on theoretical priors; second, we use "honest" machine learning approaches recently developed in the literature.

Heterogeneity along pre-specified dimensions

Table 15 analyzes heterogeneous treatment effects along five dimensions measured at baseline: proximity to the FT (for farmers only); land size; number of feeds used; herd size; ownership of crossbred cattle. These dimensions were hypothesized to be susceptible to trigger heterogeneous effects. In particular, the initial familiarity with advanced feeding practices and crossbred animals could enhance program impacts (or dampen them in case of diminishing returns to knowledge). Land and herd size proxy for assets that are needed to invest in better feeding; moreover, the program expected changes by pushing farmers with large herd and little land toward more intensive practices (fewer, better fed animals). Last, proximity to the FT can be expected to facilitate knowledge transmission and trust leading to stronger adoption.

The model in Table 15 is similar to equation (1), where the interaction variable is a single indicator.[34] We find very limited evidence of heterogeneous impacts of the FT program. Panel A shows that program impacts are very close for farmers that were reported by the FT as part of their network and for other farmers: interaction coefficients are small, and statistically

[34]For land size, feed usage and herd size, the indicator is equal to one for above median baseline values.

non significant. Similarly, Panel B of Table 15 shows that program impacts do not vary for FTs based on the size of their land; effects on profits are concentrated on farmers with more land, but the difference is only marginally statistically significant, and insignificant when taking into account multiple hypothesis testing (adjusted p-value=.14). Panel C shows more evidence of heterogeneous impacts. FTs that were using more feeds at baseline increase their usage significantly less, perhaps due to diminishing marginal returns.[35] In addition, the effects of the program on milk production and profits are concentrated among FTs and farmers with initially low levels of technology adoption: in fact, the estimates are close to 0 for those in the initial upper half.[36] This may help explain why FTs on average do not display larger program impacts than other farmers: even though they have more direct access to information through the program, they also start from higher adoption levels. Panel D finds no differences across initial herd sizes, consistent with the findings on land size. Last, there is also no evidence that experience with crossbred animals affects program impacts (Panel E).

To sum up, there is some evidence of larger program impacts for farmers and FTs starting from low adoption level. The evidence is robust to statistical adjustments for multiple hypothesis testing, and differences are quantitatively large: impacts on milk production and profits are driven by farmers with initially less advanced feeding practices. However, we do not detect any difference in other dimensions on which they could be expected.[37]

Machine learning

Recent developments in machine learning provide unstructured, yet "honest" methods to uncover heterogeneous treatment effects in contexts with wide data (rich baseline information).[38] We follow the "generic machine learning" approach proposed by Chernozhukov et al. (2018), including their use of variational estimation and inference methods to account for estimation and splitting uncertainty.

As noted by Heller & Davis (2017) in an early application of such methods, even with a sample of about 3,000 observations (which is relatively large for a randomized controlled trial with primary data collection), results may be sensitive to the selection of heterogeneity dimensions and the pre-processing of the corresponding variables. We experimented with different sets of variables, including large ones, and present here the results obtained with a smaller list of 16 variables, denoted by a star in Table 2. Table 16 shows estimates of the

[35] The difference is however not statistically significant when adjusting p-values for multiple hypothesis testing.

[36] The interaction effects are not statistically significant for FTs, but the point estimates are very close to those of other farmers.

[37] Note that the adjustments for multiple hypothesis testing conducted in this section are mildly conservative, as we only adjust for the fact that four outcomes variables are analyzed. If we were to consider families of twenty regressions (= 4×5 interaction dimensions), adjustments would be much more stringent, reinforcing the conclusion that heterogeneous effects are not large enough to be statistically detected.

[38] See in particular Athey et al. (2018) and Chernozhukov et al. (2018).

best linear predictor of conditional average treatment effects (equation 2.1 in Chernozhukov et al. (2018)) obtained from post-processing four types of algorithms: causal forest, random forest, elastic net, and boosting. The first parameter, β_1, is the average treatment effect (ITT parameter). β_2 is the parameter of interest: as shown by Chernozhukov et al. (2018), it is equal to $Cov(s_0(Z), S(Z))/Var(S(Z))$ where $S(Z)$ is the proxy predictor of the true conditional average treatment effect (CATE), $s_0(Z)$. The null hypothesis $\beta_2 = 0$ is equivalent to the combined hypothesis that there is no treatment effect heterogeneity ($s_0(Z) = s_0$) or that this heterogeneity is not well captured by the proxy predictor $S(Z)$ (so that $Cov(s_0(Z), S(Z)) = 0$. The 90% confidence intervals reported below β_2 show that we can never reject the null, irrespective of the outcome we consider or the proxy predictor we use. This may imply that our proxy predictors are performing poorly. However, the absence of detectable heterogeneity is robust to varying the sets of predictors (Z) and tuning parameters (changing, for instance, the share of observations allocated to the auxiliary and main samples). In particular, including all the variables listed in Table 2 as predictors (rather than starred variables only) makes no meaningful difference.[39] Note that Table 15 is restricted to non-FT farmers: mixing farmers and FTs would artificially create heterogeneity as the two groups have not received the same treatment. We separately investigated treatment effect heterogeneity among FTs, and found no significant heterogeneity; this may however be largely due to insufficient sample size.

Table 17 illustrates the same finding by displaying the gap between the average treatment effects at the top 20% and bottom 20% of the treatment effect distribution (so-called sorted Group Average Treatment Effects, or GATEs). Again, the difference between the two coefficients ($\gamma_5 - \gamma_1$) is not different from 0, irrespective of the outcome considered and the algorithm used. Confidence intervals are quite wide, underscoring the cost (in terms of precision) of putting so little structure on the data. Yet, the absence of statistically significant heterogeneity is informative, by comparison with other studies that find significant differences with similar sample sizes – for instance Crépon et al. (2015) as revisited in Chernozhukov et al. (2018).[40]

To sum up, the theory-driven and data-driven approaches provide limited evidence of heterogeneous treatment effects. There are two obvious caveats: limited statistical power and the fact some potentially relevant dimensions of heterogeneity may not be captured by baseline covariates – for instance, entrepreneurship skills or motivation are not directly measured. Yet these results add to the absence of detectable impact of the needs assessment variation.

[39] Note that the only variable yielding heterogeneous impact in Table 15 ("Above median baseline adoption level") is introduced as a continuous predictor in the machine learning analysis (baseline adoption index). The fact that machine learning algorithms fail to detect heterogeneous treatment effects detected by parsimonious OLS model may be an illustration of the statistical costs of putting so little structure in machine learning models.

[40] Given the failure to detect statistically significant heterogeneity, we do not perform a classification analysis, whose interpretation would remain unclear.

Overall these results hence suggest either that there is limited heterogeneity in this context, or that the core FT program - possibly by implicitly drawing on local information of the FTs about farmers in his village - already manages to account for the relevant heterogeneity. The only clear indication of heterogeneous impact suggests that the program was more effective in reaching out to farmers and FTs initially lagging behind, which possibly points in this direction. Interestingly, this is partly consistent with the intuition of ICRAF staff at the start of the intervention.

6 Conclusion

The FT program was designed as a low cost strategy to reach out to a large number of farmers, with an average cost of 160 USD per village per year. We find robust positive impacts on knowledge, adoption, and milk production, and overall positive impacts on net revenues (or profits). Taking the most conservative profit estimates, we calculate a net gain of 454 USD per village per year. These computations are only indicative as the statistical precision of profit estimates is limited and profit measurement is imperfect. The additional profit gains in the linkage and signpost variation translate in a net gain of 1800 USD, resp. 1668 USD, with the 90% lower bound at 200 USD for the linkage variation (see Appendix A.2). To the extent that improvements in knowledge lead to sustained change in practices also after the end of the program, these estimates underestimate long-term gains. Overall, these results provide evidence that the contact farmer extension model can provide scalable, cost-effective extension to smallholders, at odds with the skepticism that prevailed in the past decades. Importantly, our evaluation goes beyond the proof of concept: the specific program we evaluate had already served 40,000 dairy farmers in three countries before starting the evaluation, and its main features resemble basic volunteer farmer trainer programs found elsewhere, including the absence of monetary incentives.

Further research will aim at better understanding the role of heterogeneity and different levels of customization. The results to date, however, clearly show lack of impact of the program variation meant to provide individual customized advice, and the failure to detect heterogeneous treatment effects. These findings may seem at odds with past evidence on the importance of heterogeneity. Note however that the hypothesis that heterogeneity across smallholders particularly impedes agricultural technology diffusion in Sub-Saharan Africa was initially advanced for fertilizers, as variations in unobserved soil quality may imply large variations in returns and limit what can be learnt from experience in neighboring plots (Duflo et al. 2008, Suri 2011, Tjernström 2017). It is possible that the lack of heterogeneous effects in the case of dairy farming is due to the fact that soil quality matters less to grow cattle feeds, and not at all for such actions as watering animals ad libidum.

Even so, heterogeneity in households' other factor endowments likely point to possibility of substantial heterogeneity in returns. Indeed, in the context of the intervention analyzed, ICRAF experts insisted on the fact that some of the technologies they were promoting would not make economic sense for all farmers, which justified teaching a broad set of principles and promoting a variety of feeds and practices (about 30 technologies in total). Given the experts priors, one would hence expect the program to have heterogeneous effects, unless farmers and FTs pick different technologies relevant to them, and are all able to reach the same returns. While we cannot rule out that interpretation, it is worth highlighting that such heterogeneous adoption would need to exactly compensate existing differences so as to yield homogeneous treatment effects. The only meaningful heterogeneity we detect – larger program impacts on farmers and FTs starting from low adoption – suggests a more straightforward story in which simple, uniform practices are adopted by all, and have largest impacts on those initially lagging behind.

Treatment effect heterogeneity can also come from the heterogeneity in other constraints that farmers may face. The FT program evaluated was a component of a larger project aimed at increasing the productivity of smallholder dairy farmers through improvements in breeding stock, animal health, feeding practices and the development of dairy value chains. As such the interventions aimed at improving feeding practices occurred in a context in which some of the other constraints to productivity may have been lifted by the other components of the EADD program, potentially providing a setting allowing for homogeneous treatment response. This does not take away from the relevance of the findings, as livestock interventions often purposely target these different aspects, following the recommendations of livestock research.

More advanced approaches to customized training, and further tests of the heterogeneity hypothesis in diverse settings are promising research avenues. In the meantime, the basic contact farmer model stands out as a simple but effective tool to help increase agricultural revenues of poor smallholders.

References

Aker, J. C. (2011), 'Dial "a" for agriculture: a review of information and communication technologies for agricultural extension in developing countries', *Agricultural Economics* **42**(6), 631–647.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1574-0862.2011.00545.x*

Anderson, J. R. & Feder, G. (2007), Chapter 44 agricultural extension, *in* R. Evenson & P. Pingali, eds, 'Agricultural Development: Farmers, Farm Production and Farm Markets', Vol. 3 of *Handbook of Agricultural Economics*, Elsevier, pp. 2343 – 2378.
URL: *http://www.sciencedirect.com/science/article/pii/S1574007206030441*

Anderson, M. L. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects', *Journal of the American Statistical Association* **103**(484), 1481–1495.
URL: *https://doi.org/10.1198/016214508000000841*

Ashraf, N., Bandiera, O. & Jack, B. K. (2014), 'No margin, no mission? a field experiment on incentives for public service delivery', *Journal of Public Economics* **120**, 1 – 17.
URL: *http://www.sciencedirect.com/science/article/pii/S0047272714001546*

Assunção, J., Bragança, A. & Hemsley, P. (2019), 'Geographic heterogeneity and technology adoption: Evidence from brazil', *Land Economics* **95**(4), 599–616.

Athey, S. & Imbens, G. (2017), Chapter 3 - the econometrics of randomized experimentsa, *in* A. V. Banerjee & E. Duflo, eds, 'Handbook of Field Experiments', Vol. 1 of *Handbook of Economic Field Experiments*, North-Holland, pp. 73 – 140.
URL: *http://www.sciencedirect.com/science/article/pii/S2214658X16300174*

Athey, S., Tibshirani, J. & Wager, S. (2018), 'Generalized random forests', *Annals of Statistics (forthcoming)* .

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I. & Sulaiman, M. (2017), 'Labor Markets and Poverty in Village Economies', *Quarterly Journal of Economics* **132**(2), 811–870.
URL: *https://academic.oup.com/qje/article-abstract/132/2/811/3075123?redirectedFrom=fulltext*

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B. & Udry, C. (2015), 'A multifaceted program causes lasting progress for the very poor: evidence from six countries', *Science* **348**(6236), 1260799.
URL: *https://science.sciencemag.org/content/348/6236/1260799*

Beaman, L., BenYishay, A., Magruder, J. & Mobarak, A. M. (2018), Can network theory-based targeting increase technology adoption?, Working Paper 24912, National Bureau of

Economic Research.
URL: *http://www.nber.org/papers/w24912*

Beaman, L. & Dillon, A. (2018), 'Diffusion of agricultural information within social networks: Evidence on gender inequalities from mali', *Journal of Development of Economics* **133**, 147–161.

BenYishay, A., Jones, M., Kondylis, F. & Mobarak, A. M. (2020), 'Gender gaps in technology diffusion', *Journal of Development Economics* **143**.
URL: *https://www.sciencedirect.com/science/article/pii/S030438781930197X*

BenYishay, A. & Mobarak, A. M. (2018), 'Social Learning and Incentives for Experimentation and Communication', *The Review of Economic Studies* **86**(3), 976–1009.
URL: *https://doi.org/10.1093/restud/rdy039*

Casaburi, L., Kremer, M., Mullainathan, S. & Ramrattan, R. (2019), 'Harnessing ict to increase agricultural production: Evidence from kenya', *Working Paper* .

Chernozhukov, V., Demirer, M., Duflo, E. & Fernandez-Val, I. (2018), Generic machine learning inference on heterogenous treatment effects in randomized experiments.

Cole, S. A. & Fernando, A. N. (2020), "mobile'izing agricultural advice: Technology adoption, diffusion and sustainability', *The Economic Journal* **ueaa084,**.
URL: *https://doi.org/10.1093/ej/ueaa084*

Corral, C., Giné, Xavier Mahajan, A. & Seira, E. (2020), 'Conditionality and specificity in fertilizer recommendations: Evidence from mexico', *mimeo UC Berkeley* .

Crépon, B., Devoto, F., Duflo, E. & Parienté, W. (2015), 'Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco', *American Economic Journal: Applied Economics* **7**(1), 123–50.
URL: *http://www.aeaweb.org/articles?id=10.1257/app.20130535*

de Janvry, A. & Sadoulet, E. (2019), Transforming developing country agriculture: Removing adoption constraints and promoting inclusive value chain development, Working Papers hal-02287668, HAL.
URL: *https://ideas.repec.org/p/hal/wpaper/hal-02287668.html*

de Janvry, A., Sadoulet, E. & Tavneet, S. (2017), Field experiments in developing country agriculture, *in* A. V. Banerjee & E. Duflo, eds, 'Handbook of Economic Field Experiments', Vol. 2 of *Handbook of Economic Field Experiments*, North-Holland, pp. 427 – 466.
URL: *https://www.sciencedirect.com/science/article/pii/S2214658X16300046*

Duflo, E. (2018), Machinistas meet randomistas : useful ML tools for empirical researchers.

Duflo, E., Kremer, M. & Robinson, J. (2008), 'How high are rates of return to fertilizer? evidence from field experiments in kenya', *American Economic Review* **98**(2), 482–88.
URL: *http://www.aeaweb.org/articles?id=10.1257/aer.98.2.482*

Emerick, K. (2018), 'Trading frictions in indian village economies', *Journal of Development Economics* **132**(C), 32–56.
URL: *https://www.sciencedirect.com/science/article/pii/S0304387817301293*

Fabregas, R., Kremer, M., Schilbach, F. & Robinson, J. (2018), 'The effectiveness of public agricultural extension: Evidence from two approaches in kenya', *Working Paper* .

FAO, GDP & IFCN (2018), *Dairy Development's Impact on Poverty Reduction*, Chicago, Illinois, USA.
URL: *http://www.fao.org/3/CA0289EN/ca0289en.pdf*

Goyal, A. & Nash, J. (2017), *Reaping Richer Returns: Public Spending Priorities for African Agricultural Productivity Growth*, number 25996 *in* 'World Bank Publications', The World Bank.
URL: *http://hdl.handle.net/10986/25996*

Hanna, R., Mullainathan, S. & Schwartzstein, J. (2014), 'Learning Through Noticing: Theory and Evidence from a Field Experiment', *The Quarterly Journal of Economics* **129**(3), 1311–1353.
URL: *https://doi.org/10.1093/qje/qju015*

Havlik, P., Valin, H., Herrero, M., Obersteiner, M., Schmid, E., Rufino, M. C., Mosnier, A., Thornton, P. K., Böttcher, H., Conant, R. T., Frank, S., Fritz, S., Fuss, S., Kraxner, F. & Notenbaert, A. (2014), 'Climate change mitigation through livestock system transitions', *Proceedings of The National Academy of Science* **111**(10), 3709–3714.
URL: *https://www.pnas.org/content/111/10/3709*

Heller, S. B. & Davis, J. M. (2017), 'Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs', *American Economic Review: Papers & Proceedings* **107**(5), 546–550.

Kondylis, F., Mueller, V. & Zhu, J. (2017), 'Seeing is believing? evidence from an extension network experiment', *Journal of Development Economics* **125**, 1 – 20.
URL: *http://www.sciencedirect.com/science/article/pii/S0304387816300827*

Magruder, J. R. (2018), 'An assessment of experimental evidence on agricultural technology adoption in developing countries', *Annual Review of Resource Economics* **10**(1), 299–316.
URL: *https://doi.org/10.1146/annurev-resource-100517-023202*

Marenya, P. P. & Barrett, C. B. (2009), 'Soil quality and fertilizer use rates among smallholder farmers in western kenya', *Agricultural Economics* **40**(5), 561–572.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1574-0862.2009.00398.x*

Miguel, E. & Kremer, M. (2004), 'Worms: Identifying impacts on education and health in the presence of treatment externalities', *Econometrica* **72**(1), 159–217.
URL: *http://www.jstor.org/stable/3598853*

Munshi, K. (2004), 'Social learning in a heterogeneous population: technology diffusion in the indian green revolution', *Journal of Development Economics* **73**(1), 185 – 213.
URL: *http://www.sciencedirect.com/science/article/pii/S0304387803001342*

Muralidharan, K., Romero, M. & Wüthrich, K. (2019), Factorial designs, model selection, and (incorrect) inference in randomized experiments, Working Paper 26562, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w26562*

Romano, J. P. & Wolf, M. (2016), 'Efficient computation of adjusted p-values for resampling-based stepdown multiple testing', *Statistics and Probability Letters* **113**(C), 38–40.
URL: *https://EconPapers.repec.org/RePEc:eee:stapro:v:113:y:2016:i:c:p:38-40*

Stevenson, J., Vanlauwe, B., Macours, K., Johnson, N., Krishnan, L., Place, F., Spielman, D., Hughes, K. & Vlek, P. (2019), 'Farmer adoption of plot-and farm-level natural resource management practices: Between rhetoric and reality', *Global Food Security* **20**.

Suri, T. (2011), 'Selection and comparative advantage in technology adoption', *Econometrica* **79**(1), 159–209.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7749*

Tjernström, E. (2017), Learning from others in heterogeneous environments, Working paper.

Figure 1: Treatment and control parishes, Kamuli district

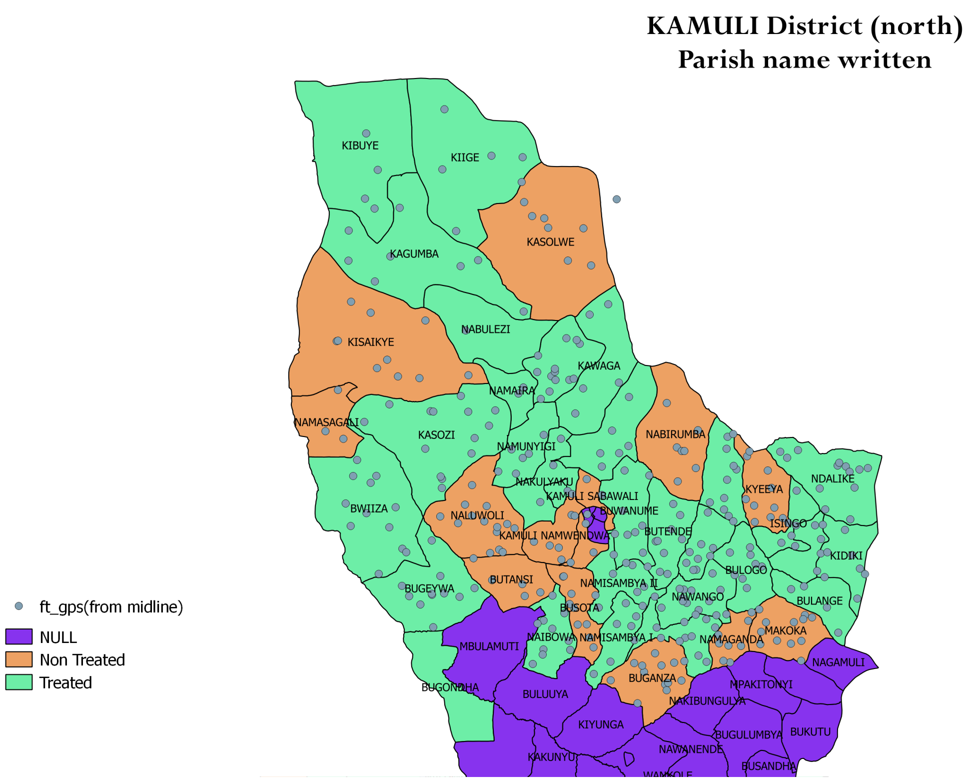

Table 1: Sample composition

	FTs		Farmers	
	N	%	N	%
	A. Whole Sample			
Untreated	216	34.45	859	34.43
Treated	411	65.55	1636	65.57
Total	627	100.00	2495	100.00
	B. Treated Sample			
By treatment variation				
Linkage variation	206	50.12	821	50.18
Needs assessment variation	203	49.39	808	49.39
Signpost variation	204	49.64	814	49.76
By number of treatment variations				
Basic treatment only	50	12.17	198	12.10
1 variation	162	39.42	644	39.36
2 variations	146	35.52	583	35.64
3 variations	53	12.90	211	12.90

Notes: This table shows the sample composition according to treatment and control groups for FTs and Farmers. Panel A shows the number of observations for the treated and the untreated as well as their corresponding percentage from the total number of FTs and Farmers. Panel B shows the number of observations and the percentage by treatment variation and by the number of treatment variations assigned.
Source: Endline survey and lottery results.

Table 2: Sample description

	Farmer trainers			Other farmers		
	Mean	Std deviation	Observations	Mean	Std deviation	Observations
A. Household head background						
Age*	42.49	11.87	603	46.46	13.95	2443
Male	0.95	0.21	603	0.91	0.29	2445
Education index (schooling + read + write)*	0.57	0.62	603	-0.14	1.03	2445
Days of illness in past 30 days	0.72	3.81	603	1.18	4.93	2445
=1 if permanent disability	0.02	0.15	603	0.04	0.19	2445
B. Milk production						
Milk production index*	0.06	1.14	604	-0.02	0.96	2458
Technology usage index*	0.54	0.96	606	-0.14	0.96	2460
Number of commercial feeds used	1.01	0.81	606	0.87	0.68	2460
Technology knowledge index*	0.57	0.93	606	-0.14	0.97	2460
Years of experience with cross-bred cattle	2.05	4.05	607	1.69	4.00	2460
Years of experience with indigenous cattle	10.17	9.73	606	12.47	11.48	2459
C. Cattle						
# crossbred cows*	0.62	2.48	608	0.40	1.24	2459
# indigenous cows*	1.58	2.76	608	1.97	3.81	2459
# crossbred heads of cattle*	1.69	6.85	608	1.10	3.14	2459

Table 2: (continued)

	Farmer trainers			Other farmers		
	Mean	Std deviation	Observations	Mean	Std deviation	Observations
# indigenous heads of cattle*	4.50	6.65	608	5.40	8.96	2459
2nd quartile cows per acre	0.26	0.44	608	0.25	0.44	2459
3rd quartile cows per acre	0.21	0.41	608	0.26	0.44	2459
4th quartile cows per acre	0.27	0.45	608	0.24	0.43	2459
=1 if no pasture	0.25	0.43	608	0.22	0.41	2459
=1 if cattle is vaccinated	0.13	0.33	608	0.14	0.34	2459
=1 if used artificial insemination or controlled breeding	0.22	0.41	608	0.07	0.25	2459
D. Pasture						
Pasture land size (acres)*	2.20	5.06	607	2.21	5.47	2449
=1 if water source on pasture*	0.33	0.47	607	0.36	0.48	2459
E. Capital						
# assets held on list 24*	6.71	1.94	608	5.63	1.84	2459
Land size (in acres)*	8.84	15.29	604	8.15	13.37	2444
Cultivation land size (in acres)*	5.05	5.73	605	4.00	3.94	2446
F. Labor						
Total hours worked with livestock in a day*	12.89	12.00	608	10.66	8.92	2459

Table 2: (continued)

	Farmer trainers			Other farmers		
	Mean	Std deviation	Observations	Mean	Std deviation	Observations
G. Social capital						
=1 if in FT network*	-	-	-	0.75	0.43	2459
=1 if one member of HH is village committee member	0.33	0.47	608	0.21	0.41	2459
=1 if household head belongs to a farmer group	0.19	0.40	604	0.15	0.35	2444
H. FT characteristics						
=1 if has a male FT	0.95	0.22	596	0.95	0.22	2412
=1 if FT has completed high school	0.34	0.47	596	0.34	0.47	2412
=1 if FT and household head are of same gender	0.88	0.33	608	0.84	0.36	2459

Notes. For each baseline variable listed in the left column, the table displays the means and standard deviations among farmer trainers and other farmers. Indexes are computed as weighted sums of Z-scores, following Anderson (2008). Variables with a star are used as predictor of endline outcomes by the machine learning algorithms used in Section 5.

Table 3: FT program impact on main outcomes

	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N
	Farmer trainers					Other farmers				
Outcome (standardized index)										
Knowledge	0.48***	(0.06)	0.00	0.00	607	0.08*	(0.04)	0.07	0.22	2459
Adoption	0.19***	(0.04)	0.00	0.00	607	0.06**	(0.03)	0.02	0.06	2460
Milk production	0.16	(0.13)	0.21	0.40	606	0.13**	(0.05)	0.01	0.06	2459
Profits	0.05	(0.09)	0.59	0.63	607	0.07	(0.05)	0.16	0.22	2460

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Outcomes are standardized indexes, computed as the weighted sum of Z-scores of individual variables following Anderson (2008); they are normalized to have a standard deviation equal to 1 in the control group. Each regression includes strata fixed effects, and controls for baseline value of the outcome following equation (1) in the text. Standard errors are clustered at the parish level. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of the four indexes.
Source. Endline and baseline surveys.

Table 4: Impact of variations on main outcomes

	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N
	Farmer trainers					Other farmers				
Outcome (standardized index)	A. Needs assessment									
Knowledge	0.32***	(0.11)	0.00	0.03	395	0.05	(0.05)	0.26	0.75	1596
Adoption	-0.02	(0.05)	0.70	0.97	395	0.03	(0.02)	0.18	0.60	1596
Milk production	-0.01	(0.2)	0.94	0.97	395	0.01	(0.06)	0.93	0.75	1595
Profits	-0.14	(0.11)	0.19	0.41	395	0.02	(0.06)	0.73	0.75	1596
	B. Linkage									
Knowledge	-0.13	(0.11)	0.24	0.58	395	0.07	(0.05)	0.17	0.14	1596
Adoption	0.03	(0.05)	0.56	0.65	395	0.04*	(0.02)	0.09	0.12	1596
Milk production	0.12	(0.2)	0.56	0.73	395	0.17**	(0.06)	0.01	0.02	1595
Profits	0.06	(0.11)	0.57	0.73	395	0.15**	(0.06)	0.01	0.04	1596
	C. Signpost									
Knowledge	0.05	(0.11)	0.62	0.90	395	-0.03	(0.05)	0.55	0.54	1596
Adoption	0.04	(0.05)	0.43	0.90	395	0.05**	(0.02)	0.04	0.11	1596
Milk production	-0.18	(0.2)	0.36	0.79	395	0.11*	(0.06)	0.08	0.32	1595
Profits	0.06	(0.11)	0.58	0.90	395	0.14**	(0.06)	0.02	0.22	1596

Notes. Each line shows the estimated effect of winning the lottery giving access to one of the three variations (corresponding to the three panels) on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Outcomes are standardized indexes, computed as the weighted sum of Z-scores of individual variables following Anderson (2008); they are normalized to have a standard deviation equal to 1 in the control group. Each regression includes strata fixed effects, and controls for baseline value of the outcome following equation (1) in the text. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of the four indexes.

Source. Endline and baseline surveys.

Table 5: FT program impact on observed cow size, health, and milk quality

			Basic specification		With agrovet fixed effects	
	N	Mean Control	β	Std. Error	β	Std. Error
Panel A: Impact of assignment to farmer trainer program: FT and farmers together						
Cow size and health						
average cow size (circumference in cm)	938	152.9	1.522**	(0.70)	1.409**	(0.69)
average index of bad health (first factor)	938	0.017	-0.008	(0.074)	0.024	(0.054)
Milk quality						
fat content	901	4.519	0.007	(0.18)	n.a.	
standardized fat content (accounting for breed and session)	901	-0.026	0.017	(0.086)	n.a.	
standardized fat content (accounting for breed and session) - excluding quality flags	850	-0.017	0.011	(0.086)	n.a.	
Panel B: Impact of assignment to farmer trainer program: FT						
Cow size and health						
average cow size (circumference in cm)	420	154.6	0.856	(0.98)	0.876	(0.94)
average index of bad health (first factor)	420	-0.044	0.104	(0.089)	0.142**	(0.067)
Milk quality						
fat content	489	4.556	-0.068	(0.23)	n.a.	
standardized fat content (accounting for breed and session)	489	0.013	-0.016	(0.11)	n.a.	
standardized fat content (accounting for breed and session) - excluding quality flags	464	0.008	-0.005	(0.11)	n.a.	
Panel C: Impact of assignment to farmer trainer program:farmers						
Cow size and health						
average cow size (circumference in cm)	518	151.6	2.144**	(0.96)	1.958**	(0.96)
average index of bad health (first factor)	518	0.066	-0.0957	(0.084)	-0.0600	(0.069)
Milk quality						
fat content	412	4.479	0.065	(0.20)	n.a.	
standardized fat content (accounting for breed and session)	412	-0.068	0.042	(0.097)	n.a.	
standardized fat content (accounting for breed and session) - excluding quality flags	386	-0.043	0.014	(0.10)	n.a.	

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program on the outcome listed in the left column. Each regression includes strata fixed effects, and specification in the last two columns include enumerator (agrovets) fixed effects.
Source. Unannounced visits to farms after endline survey.

Table 6: Impact on dairy profits

	Mean Control	T	Std. Error	p-value	*N*	Mean Control	T	Std. Error	p-value	*N*
	Panel A: Impact of assignment to farmer trainer program									
	FTs					Farmers				
Detail on annual revenues and annual expenditures										
Annual revenues from milk production and other earnings	2362.70	150.63	(205.85)	0.47	607	1913.99	129.48	(112.41)	0.25	2460
Milk production	1438.81	156.96	(161.25)	0.33	607	1157.83	135.00	(80.61)	0.10	2460
Livestock sales	860.57	-28.16	(103.17)	0.79	607	702.74	-16.17	(51.14)	0.75	2460
Earnings from manure/oxen/fees for impregnation	63.31	21.83	(17.56)	0.22	607	53.43	10.65	(13.64)	0.44	2460
Annual expenditures	618.24	188.96*	(99.68)	0.06	607	460.86	47.72	(34.15)	0.17	2460
Expenditures: Feeds/feeding practices	87.60	87.90**	(32.77)	0.01	607	55.32	5.36	(7.99)	0.50	2460
Expenditures: Seeds	0.95	1.29**	(0.53)	0.02	607	1.17	-0.24	(0.35)	0.50	2460
Expenditures: Health	236.05	-10.04	(32.81)	0.76	607	182.50	1.38	(11.55)	0.90	2460
Expenditures: Livestock acquisition	144.93	99.14**	(43.05)	0.02	607	112.34	48.89**	(17.64)	0.01	2460
Expenditures: Hired labor	86.39	14.99	(25.13)	0.55	607	64.07	-2.79	(11.42)	0.81	2460
Expenditures: Fencing/paddocking/barns/storage	30.63	0.73	(11.44)	0.95	607	16.38	1.10	(5.54)	0.84	2460
Expenditures: Extension advice	0.86	-0.41	(0.56)	0.47	607	0.12	0.10	(0.08)	0.20	2460
Expenditures: Compensation for animal destruction	31.78	-3.36	(5.72)	0.56	607	30.14	-6.31**	(2.98)	0.04	2460
Annual profits										
Excluding livestock sales/acquisition and feed expenses	1116.42	176.87	(140.13)	0.21	607	918.05	152.18*	(74.87)	0.05	2460
With livestock sales/acquisition but excluding feed expenses	1028.82	88.97	(135.46)	0.51	607	862.72	146.82*	(73.09)	0.05	2460
With livestock sales/acquisition and feed expenses	1744.46	-38.33	(153.04)	0.8	607	1453.12	81.76	(102.01)	0.43	2460

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Each regression includes strata fixed effects. Standard errors are clustered at the parish level. Annual profits in the last line are computed as Annual revenues minus Annual expenditures. The two lines above exclude noisy elements from the computation (feed expenses, and net investment in livestock).

Source. Endline survey.

Table 7: Technical information

Outcome	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N
	Farmer trainers						Other farmers					
A. Main treatment												
Received advice on feeds/feeding practices from EA/Vet.	0.08	0.04*	(0.02)	0.09	0.11	607	0.02	0.01*	(0.01)	0.06	0.07	2460
Number of feeds/feeding practices for which received advice from EA/Vet.	0.18	0.50***	(0.16)	0.00	0.03	607	0.04	0.06**	(0.03)	0.02	0.06	2460
B. Needs assessment												
Received advice on feeds/feeding practices from EA/Vet.	0.11	0.02	(0.03)	0.55	0.79	395	0.02	0.01	(0.01)	0.16	0.26	1596
Number of feeds/feeding practices for which received advice from EA/Vet.	0.58	0.09	(0.32)	0.77	0.79	395	0.09	0.01	(0.04)	0.83	0.82	1596
C. Linkage												
Received advice on feeds/feeding practices from EA/Vet.	0.1	0.03	(0.03)	0.34	0.33	395	0.03	0.01	(0.01)	0.36	0.54	1596
Number of feeds/feeding practices for which received advice from EA/Vet.	0.35	0.55*	(0.32)	0.08	0.15	395	0.09	0.01	(0.04)	0.88	0.89	1596
D. Signpost												
Received advice on feeds/feeding practices from EA/Vet.	0.1	0.04	(0.03)	0.23	0.36	395	0.03	0.00	(0.01)	0.67	0.66	1596
Number of feeds/feeding practices for which received advice from EA/Vet.	0.48	0.28	(0.32)	0.37	0.36	395	0.07	0.05	(0.04)	0.23	0.39	1596

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program (Panel A) or to the variation corresponding to Panels B to D on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Each regression includes strata fixed effects. In Panel A, standard errors are clustered at the parish level. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of two outcomes.

Source. Endline survey.

Table 8: Incentives

Outcome	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N
	Farmer trainers						Other farmers					
A. Main treatment												
Importance of helping others as driving effort	0	4.63***	(0.12)	0.00	0.00	607	0	2.29***	(0.11)	0.00	0.00	2460
Pro-social motivation (reported by FT, standardized)	-0.09	0.16**	(0.07)	0.02	0.02	607	0	-	-	-	-	-
Advice given by FT to other farmers is helpful	0	-	-	-	-	-	0.61	0.12***	(0.03)	0.00	0.00	2460
FT has given individual advice suited for each farm	0.25	0.45***	(0.04)	0.00	0.00	607	0.1	0.12***	(0.02)	0.00	0.00	2460
EA participated in a training let by FT	0.01	0.12***	(0.02)	0.00	0.00	607	0.07	0.07***	(0.01)	0.00	0.00	2460
B. Needs assessment												
Importance of helping others as driving effort		0.57**	0.24	0.02	0.07	395		0.26	0.16	0.10	0.28	1596
Pro-social motivation (reported by FT)		0.06	0.08	0.46	0.72	395		-	-	-	-	-
Advice given by FT to other farmers is helpful		-	-	-	-	-		0.01	0.02	0.78	0.95	1596
FT has given individual advice suited for each farm		0.00	0.05	0.36	0.72	395		-0.02	0.02	0.42	0.78	1596
EA participated in a training let by FT		0.00	0.03	0.93	0.91	395		0.00	0.02	0.99	0.99	1596
C. Linkage												
Importance of helping others as driving effort		0.14	0.24	0.57	0.56	395		0.24	0.16	0.13	0.26	1596
Pro-social motivation (reported by FT)		-0.12	0.08	0.14	0.35	395		-	-	-	-	-
Advice given by FT to other farmers is helpful		-	-	-	-	-		0.04*	0.02	0.06	0.17	1596
FT has given individual advice suited for each farm		0.07	0.05	0.15	0.35	395		0.02	0.02	0.24	0.26	1596
EA participated in a training let by FT		0.09**	0.03	0.01	0.02	395		0.06***	0.02	0.00	0.00	1596
D. Signpost												
Importance of helping others as driving effort		0.05	0.24	0.82	0.91	395		0.35**	0.16	0.02	0.08	1596
Pro-social motivation (reported by FT)		0.05*	0.08	0.57	0.91	395		-	-	-	-	-
Advice given by FT to other farmers is helpful		-	-	-	-	-		0.03	0.02	0.22	0.25	1596
FT has given individual advice suited for each farm		0.03*	0.05	0.45	0.91	395		0.04*	0.02	0.05	0.13	1596
EA participated in a training let by FT		0.02**	0.03	0.46	0.91	395		0.03	0.02	0.13	0.25	1596

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program (Panel A) or to the variation corresponding to Panels B to D on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Each regression includes strata fixed effects. In Panel A, standard errors are clustered at the parish level. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of five outcomes.

Source. Endline survey.

Table 9: Interactions

	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N
	Farmer trainers						Other farmers					
Outcome												
A. Main treatment												
How knowledgeable is FT regarding feeding practices?	-	-	-	-	-	-	0.49	0.18***	(0.03)	0.00	0.00	2460
How knowledgeable is FT regarding other aspects of milk farming?	-	-	-	-	-	-	0.46	0.17***	(0.03)	0.00	0.00	2460
Talked/asked advice to another villager about dairy farming?	0.51	0.26***	(0.04)	0.00	0.00	607	0.22	0.07***	(0.02)	0.00	0.00	2460
Number of people in the village talked about dairy farming	2.45	2.78***	(0.66)	0.00	0.00	607	0.52	0.12	(0.08)	0.16	0.19	2460
B. Needs assessment												
How knowledgeable is FT regarding feeding practices?	-	-	-	-	-	-	0.68	-0.02	(0.02)	0.50	0.68	1596
How knowledgeable is FT regarding other aspects of milk farming?	-	-	-	-	-	-	0.63	-0.01	(0.02)	0.62	0.68	1596
Talked/asked advice to another villager about dairy farming?	0.8	-0.04	(0.04)	0.39	0.65	395	0.31	-0.04*	(0.02)	0.07	0.18	1596
Number of people in the village talked about dairy farming	5.27	-0.05	(1.06)	0.96	0.96	395	0.66	-0.07	(0.07)	0.34	0.66	1596
C. Linkage												
How knowledgeable is FT regarding feeding practices?	-	-	-	-	-	-	0.68	-0.02	(0.02)	0.43	0.62	1596
How knowledgeable is FT regarding other aspects of milk farming?	-	-	-	-	-	-	0.62	0.01	(0.02)	0.75	0.76	1596
Talked/asked advice to another villager about dairy farming?	0.79	-0.02	(0.04)	0.70	0.71	395	0.26	0.05**	(0.02)	0.04	0.11	1596
Number of people in the village talked about dairy farming	4.43	1.58	(1.05)	0.13	0.32	395	0.53	0.18**	(0.07)	0.01	0.05	1596
D. Signpost												
How knowledgeable is FT regarding feeding practices?	-	-	-	-	-	-	0.65	0.03	(0.02)	0.15	0.36	1596
How knowledgeable is FT regarding other aspects of milk farming?	-	-	-	-	-	-	0.6	0.05**	(0.02)	0.03	0.14	1596
Talked/asked advice to another villager about dairy farming?	0.75	0.06	(0.04)	0.16	0.36	395	0.28	0.01	(0.02)	0.77	0.76	1596
Number of people in the village talked about dairy farming	4.88	0.78	(1.06)	0.46	0.52	395	0.58	0.08	(0.07)	0.27	0.44	1596

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program (Panel A) or to the variation corresponding to Panels B to D on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Each regression includes strata fixed effects. In Panel A, standard errors are clustered at the parish level. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of four outcomes.

Source. Endline survey.

Table 10: Complementary interventions

	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N	Mean control	β	Std. Error	Unadj. p-value	R-Wolf p-value	N
	Farmer trainers						Other farmers					
Outcome	A. Main treatment											
Ever received seeds from EADD	0.09	0.72***	(0.03)	0.00	0.00	607	-	-	-	-	-	-
Ever distributed seeds to other farmers	0.17	0.36***	(0.04)	0.00	0.00	607	-	-	-	-	-	-
Number of exotic cows (up to 10, report from paravet visits)	0.71	0.08	(0.11)	0.46	0.76	521	0.46	0.22**	(0.11)	0.04	0.06	548
Number of indigenous cows (up to 10, report from paravet visits)	0.97	-0.03	(0.16)	0.85	0.86	521	1.49	-0.09	(0.14)	0.54	0.56	548
	B. Needs assessment											
Ever received seeds from EADD	0.78	0.08**	(0.04)	0.03	0.10	395	-	-	-	-	-	-
Ever distributed seeds to other farmers	0.48	0.09*	(0.05)	0.06	0.18	395	-	-	-	-	-	-
Number of exotic cows (up to 10, report from paravet visits)	0.79	-0.10	(0.14)	0.48	0.73	334	0.58	0.15	(0.12)	0.20	0.35	349
Number of indigenous cows (up to 10, report from paravet visits)	1.07	-0.10	(0.18)	0.57	0.73	334	1.34	0.17	(0.2)	0.40	0.41	349
	C. Linkage											
Ever received seeds from EADD	0.76	0.11***	(0.04)	0.00	0.02	395	-	-	-	-	-	-
Ever distributed seeds to other farmers	0.51	0.03	(0.05)	0.56	0.87	395	-	-	-	-	-	-
Number of exotic cows (up to 10, report from paravet visits)	0.69	0.10	(0.14)	0.48	0.87	334	0.56	0.15	(0.12)	0.18	0.34	349
Number of indigenous cows (up to 10, report from paravet visits)	0.98	0.05	(0.18)	0.76	0.87	334	1.53	-0.23	(0.2)	0.26	0.34	349
	D. Signpost											
Ever received seeds from EADD	0.82	0.00	(0.04)	0.94	0.96	395	-	-	-	-	-	-
Ever distributed seeds to other farmers	0.51	0.03	(0.05)	0.52	0.90	395	-	-	-	-	-	-
Number of exotic cows (up to 10, report from paravet visits)	0.74	0.04	(0.14)	0.79	0.96	334	0.58	0.15	(0.12)	0.19	0.35	349
Number of indigenous cows (up to 10, report from paravet visits)	0.87	0.23	(0.18)	0.19	0.55	334	1.47	-0.11	(0.2)	0.59	0.59	349

Notes. Each line shows the estimated effect of winning the lottery giving access to the FT program (Panel A) or to the variation corresponding to Panels B to D on the outcome listed in the left column. A separate OLS regression is estimated for farmer trainers and other farmers. Each regression includes strata fixed effects. In Panel A, standard errors are clustered at the parish level. Romano-Wolf p-values are adjusted to control the family-wise error rate in the family of four outcomes.

Source. Endline survey.

Table 11: Intra-village correlation

	FTs and other farmers	Other farmers only
Education	0.02	0.08
	[0 ; 0.05]	[0.04 ; 0.11]
Knowledge	0.26	0.39
	[0.22 ; 0.3]	[0.34 ; 0.43]
Adoption	0.27	0.47
	[0.23 ; 0.3]	[0.43 ; 0.51]
Milk production	0.09	0.08
	[0.06 ; 0.12]	[0.04 ; 0.11]
# crossbred cows	0.09	0.12
	[0.06 ; 0.12]	[0.09 ; 0.16]
# indigenous cows	0.14	0.15
	[0.11 ; 0.18]	[0.11 ; 0.19]

Notes. For each variable listed in the left column, the table displays the intra-village coefficient of correlation and its 95% confidence interval, in the complete sample of farmer trainers and other farmers, or in the sample of other farmers only. Variables are a subset of those in Table 2.
Source. Baseline survey.

Table 12: More intra-village correlations

	FTs and other farmers	Other farmers only
Years of experience with cross-bred cattle	0.17	0.17
	[0.14 ; 0.21]	[0.13 ; 0.21]
Years of experience with indigenous cattle	0.07	0.07
	[0.04 ; 0.1]	[0.03 ; 0.11]
# crossbred heads of cattle*	0.10	0.16
	[0.07 ; 0.13]	[0.12 ; 0.2]
# indigenous heads of cattle*	0.18	0.19
	[0.15 ; 0.22]	[0.15 ; 0.23]
# cows per acre of land	0.09	0.11
	[0.06 ; 0.12]	[0.07 ; 0.15]
Pasture land size (acres)*	0.14	0.17
	[0.11 ; 0.18]	[0.12 ; 0.21]
# assets held on list 24*	0.09	0.12
	[0.06 ; 0.12]	[0.09 ; 0.16]
Land size (in acres)*	0.12	0.13
	[0.09 ; 0.16]	[0.09 ; 0.17]
Cultivation land size (in acres)*	0.08	0.11
	[0.05 ; 0.11]	[0.07 ; 0.15]
Total hours worked with livestock in a day*	0.23	0.29
	[0.2 ; 0.27]	[0.25 ; 0.33]

Notes. For each variable listed in the left column, the table displays the intra-village coefficient of correlation and its 95% confidence interval, in the complete sample of farmer trainers and other farmers, or in the sample of other farmers only. Variables are a subset of those in Table 2.
Source. Baseline survey.

Table 13: Spillover analysis

	Knowledge			Adoption			Milk production			Profits		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.074*	0.171**	0.082	0.058**	0.111**	0.091**	0.147**	0.263**	0.150**	0.090	0.120	0.090
	(0.041)	(0.083)	(0.062)	(0.027)	(0.050)	(0.038)	(0.059)	(0.113)	(0.073)	(0.059)	(0.109)	(0.065)
NT_{02}		0.014			-0.008			-0.014			0.000	
		(0.045)			(0.022)			(0.033)			(0.034)	
NT_{25}		0.004			0.011**			0.014*			0.014	
		(0.009)			(0.005)			(0.007)			(0.009)	
$NT_{02} \times T$		-0.029			0.003			-0.039			-0.019	
		(0.039)			(0.022)			(0.027)			(0.034)	
$NT_{25} \times T$		-0.002			-0.005			0.005			0.003	
		(0.007)			(0.004)			(0.007)			(0.007)	
$T_{02} \times (1-T)$			0.016			0.065			0.006			0.000
			(0.100)			(0.048)			(0.057)			(0.056)
Observations	2,459	2,459	2,459	2,460	2,460	2,460	2,459	2,459	2,459	2,460	2,460	2,460
Mean outcome in control	-0.103	-0.103	-0.103	-0.088	-0.088	-0.088	-0.055	-0.055	-0.055	-0.038	-0.038	-0.038
Implied total effect	0.049	0.067	0.056	0.038	0.136	0.071	0.096	0.246	0.099	0.059	0.202	0.059
se	0.027	0.108	0.055	0.017	0.062	0.032	0.039	0.085	0.054	0.038	0.102	0.047
p-value total effet	0.076	0.537	0.306	0.034	0.031	0.030	0.015	0.005	0.072	0.131	0.051	0.217
Implied spillover effect		-0.045	0.003		0.064	0.011		0.074	0.001		0.123	0.000
se		0.106	0.018		0.056	0.008		0.073	0.010		0.108	0.010
p-value spillover effect		0.674	0.877		0.256	0.184		0.312	0.915		0.257	0.994

Notes. Each column displays a separate regression. Model 1 corresponds to equation (1) in the text (without baseline covariates) and does not account for spillovers across villages. Models 2 and 3 correspond to equations (3) and (4) in the text to take into account these spillovers. The standardized outcomes considered are the same as in Table 3. Farmer trainers are excluded. The "implied total effect" is the weighted average of effects on control and treated farmers, combining the direct effect of the treatment and, in models 2 and 3, spillover effects computed at mean exposure. The "implied spillover effect" is the weighted average of spillover effects on control and treated farmers. T is the indicator variable for winning the lottery giving the village access to the program.NT_{02} (resp. NT_{25} is the number of (treated) FTs living between 0 and 2 kms (resp. 2 and 5 kms) from the farmer. T_{02} is an indicator variable equal to 1 if the farmer has at least one treated farmer living less than 2 kms away from him.
Source. Endline survey.

Table 14: Heterogeneity by FT network membership and program participation

	(1)	(2)	(3)	(4)	(5)	(6)
Predictor (baseline standardized index / distance to FT's index)						
	A. Member of FT network					
Knowledge	0.014	0.015			0.019	0.017
	(0.013)	(0.014)			(0.016)	(0.017)
Adoption	0.005	0.008			0.002	0.006
	(0.012)	(0.013)			(0.013)	(0.014)
Milk production	0.035***	0.013			0.047***	0.020
	(0.010)	(0.017)			(0.012)	(0.021)
Knowledge distance			-0.009	-0.005	0.003	0.005
			(0.014)	(0.015)	(0.017)	(0.017)
Adoption distance			-0.013	-0.013	-0.009	-0.009
			(0.012)	(0.013)	(0.013)	(0.013)
Milk production distance			0.006	0.005	-0.016*	-0.011
			(0.008)	(0.013)	(0.010)	(0.017)
Other controls		x		x		x
Observations	2,492	2,457	2,480	2,417	2,480	2,417
Ajdusted R^2	0.005	0.011	-0.000	-0.000	0.005	0.007
	B. Trainee of FT					
Knowledge	-0.002	-0.000			-0.001	-0.003
	(0.015)	(0.016)			(0.018)	(0.019)
Adoption	0.044***	0.033**			0.046***	0.037**
	(0.013)	(0.014)			(0.014)	(0.015)
Milk production	-0.007	0.013			0.008	0.017
	(0.011)	(0.019)			(0.014)	(0.023)
Knowledge distance			-0.007	0.003	-0.001	0.001
			(0.016)	(0.016)	(0.018)	(0.019)
Adoption distance			-0.008	-0.010	0.009	0.007
			(0.014)	(0.014)	(0.014)	(0.015)
Milk production distance			-0.012	-0.003	-0.019*	-0.006
			(0.009)	(0.015)	(0.011)	(0.018)
Other controls		x		x		x
Observations	2,492	2,457	2,480	2,417	2,480	2,417
Ajdusted R^2	0.004	0.021	-0.000	0.014	0.004	0.023
	C. FT status (benchmark)					
Knowledge	0.107***	0.082***				
	(0.011)	(0.011)				
Adoption	0.084***	0.080***				
	(0.009)	(0.009)				
Milk production	-0.002	-0.032***				
	(0.008)	(0.012)				
Other controls		x				
Observations	3,116	3,074				
Ajdusted R^2	0.103	0.216				

Notes. Each column displays the coefficients of an OLS regression The outcome is the indicator variable equal to 1 if the farmer is a member of the FT network (Panel A), a trainee of the FT after the program started (Panel B), or the FT himself (Panel C). The regressors are the baseline variables listed in the left column and include the variables of Table 2. In columns (2), (4) and (6), those controls are expressed in terms of distance to the FT (see text).

Source. Baseline and endline surveys.

Table 15: FT program impact on main outcomes: Heterogeneity analysis

	Main Effect		Interaction Term			Main Effect		Interaction Term		
	T	p-value	T	p-value		T	p-value	T	p-value	
	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*
	Farmer trainers					Other farmers				
Outcome (standardized index)	A. Farmer is a member of the FT network									
Knowledge	-	-	-	-	-	0.08	0.32	0.00	0.98	2459
						(0.08)	[0.65]	(0.09)	[0.99]	
Adoption	-	-	-	-	-	0.02	0.54	0.05	0.26	2460
						(0.04)	[0.65]	(0.04)	[0.64]	
Milk production	-	-	-	-	-	0.10	0.27	0.06	0.52	2459
						(0.09)	[0.65]	(0.1)	[0.88]	
Profits	-	-	-	-	-	0.10	0.33	-0.01	0.92	2460
						(0.1)	[0.65]	(0.1)	[0.99]	
	B. Above median baseline land size									
Knowledge	0.57***	0.00	-0.23	0.19	607	0.10*	0.07	-0.07	0.32	2459
	(0.1)	[0]	(0.17)	[0.49]		(0.06)	[0.25]	(0.07)	[0.54]	
Adoption	0.17***	0.00	0.05	0.55	607	0.04	0.26	0.04	0.36	2460
	(0.05)	[0.01]	(0.08)	[0.85]		(0.03)	[0.41]	(0.04)	[0.54]	
Milk production	0.13	0.29	0.21	0.47	606	0.07	0.16	0.13	0.30	2459
	(0.12)	[0.55]	(0.29)	[0.85]		(0.05)	[0.41]	(0.12)	[0.51]	
Profits	0.08	0.36	0.11	0.64	607	-0.03	0.58	0.22*	0.05	2460
	(0.09)	[0.55]	(0.22)	[0.85]		(0.06)	[0.59]	(0.11)	[0.14]	

Table 15: (continued)

	Main Effect		Interaction Term			Main Effect		Interaction Term		
	T	p-value	T	p-value		T	p-value	T	p-value	
	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*
	Farmer trainers					Other farmers				
					C. Above median baseline adoption level					
Knowledge	0.44***	0.00	0.13	0.46	607	0.10*	0.08	-0.04	0.60	2459
	(0.09)	[0]	(0.18)	[0.65]		(0.06)	[0.25]	(0.08)	[0.84]	
Adoption	0.26***	0.00	-0.17**	0.03	607	0.05	0.18	0.02	0.66	2460
	(0.05)	[0]	(0.08)	[0.13]		(0.04)	[0.25]	(0.05)	[0.84]	
Milk production	0.30	0.12	-0.14	0.60	606	0.23***	0.00	-0.16*	0.09	2459
	(0.19)	[0.16]	(0.26)	[0.65]		(0.07)	[0.02]	(0.1)	[0.27]	
Profits	0.24*	0.06	-0.22	0.26	607	0.21**	0.01	-0.25**	0.02	2460
	(0.12)	[0.13]	(0.2)	[0.58]		(0.08)	[0.04]	(0.11)	[0.06]	
					D. Above median baseline herd size					
Knowledge	0.61***	0.00	-0.30	0.15	607	0.04	0.46	0.08	0.37	2459
	(0.11)	[0]	(0.21)	[0.37]		(0.06)	[0.66]	(0.08)	[0.6]	
Adoption	0.23***	0.00	-0.11	0.17	607	0.07**	0.02	-0.04	0.33	2460
	(0.05)	[0]	(0.08)	[0.5]		(0.03)	[0.05]	(0.04)	[0.6]	
Milk production	0.01	0.94	0.42	0.26	606	0.06	0.27	0.21	0.14	2459
	(0.15)	[0.96]	(0.37)	[0.5]		(0.06)	[0.63]	(0.14)	[0.19]	
Profits	0.02	0.83	0.17	0.49	607	0.02	0.67	0.19	0.10	2460
	(0.1)	[0.96]	(0.25)	[0.53]		(0.05)	[0.68]	(0.11)	[0.19]	

Table 15: (continued)

	Main Effect		Interaction Term			Main Effect		Interaction Term		
	T	p-value	T	p-value		T	p-value	T	p-value	
	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*	(SE)	[adj. p-value]	(SE)	[adj. p-value]	*N*
	Farmer trainers					Other farmers				
	E. Holding crossbred cattle at baseline									
Knowledge	0.52***	0.00	-0.10	0.58	607	0.05	0.33	0.08	0.36	2459
	(0.1)	[0]	(0.18)	[0.81]		(0.05)	[0.48]	(0.08)	[0.53]	
Adoption	0.16***	0.00	0.07	0.45	607	0.05	0.11	0.01	0.72	2460
	(0.05)	[0]	(0.09)	[0.81]		(0.03)	[0.19]	(0.04)	[0.71]	
Milk production	0.03	0.72	0.49	0.13	606	0.07**	0.04	0.19	0.18	2459
	(0.08)	[0.75]	(0.32)	[0.44]		(0.03)	[0.41]	(0.14)	[0.27]	
Profits	0.09	0.31	0.12	0.60	607	0.02	0.62	0.17	0.19	2460
	(0.08)	[0.4]	(0.23)	[0.81]		(0.04)	[0.66]	(0.13)	[0.32]	

Notes. Each line displays the estimates from two separate OLS regressions, for farmers and other farmers. The outcomes are the standardized indexes of Table 3. The two coefficients correspond to the effect of winning the lottery giving access to the FT program, and to the effect of the interaction of winning the lottery with the variable in the panel heading. In Panel A, the interaction variable is an indicator equal to 1 if the farmer is a member of the FT network. In Panel B, it is an indicator equal to 1 if the farmer had above median land size at baseline. In Panel C, it is an indicator equal to 1 if the farmer had adopted more than the median number of feeds and feeding practices at baseline. In Panel D, it is an indicator equal to 1 if the farmer had above median herd at baseline. In Panel E, it is an indicator equal to 1 if the farmer held crossbred cattle at baseline.
Source. Endline and baseline surveys.

Table 16: Impact heterogeneity: best linear predictors

	Causal forest		Random forest		Elastic net		Boosting	
	ATE (β_1)	HET (β_2)	ATE (β_1)	HET (β_2)	ATE (β_1)	HET (β_2)	ATE (β_1)	HET (β_2)
Knowledge	0.081	-0.489	0.076	0.051	0.079	-0.418	0.075	-0.297
	[-0.036; 0.202]	[-2.259; 1.349]	[-0.042; 0.195]	[-0.313; 0.404]	[-0.040; 0.196]	[-1.325; 0.541]	[0.039; 0.193]	[-0.857; 0.218]
Adoption	0.056	0.981	0.05	0.12	0.052	0.195	0.054	0.275
	[-0.027; 0.123]	[-0.611; 2.645]	[-0.021; 0.121]	[-0.270; 0.532]	[-0.022; 0.125]	[-0.249; 0.807]	[-0.004; 0.114]	[-0.157; 0.679]
Milk production	0.119	0.914	0.104	0.337	0.078	0.128	0.113	0.177
	[0.004; 0.218]	[-1.232; 2.990]	[0.007; 0.214]	[-0.603; 1.325]	[-0.015; 0.175]	[-0.409; 0.992]	[0.001; 0.226]	[-0.709; 0.961]
Profits	0.081	1.273	0.078	0.301	0.056	0.256	0.083	0.36
	[-0.030; 0.191]	[-0.309; 2.989]	[-0.021; 0.187]	[-0.457; 1.028]	[-0.060; 0.151]	[-0.402; 0.898]	[-0.027; 0.194]	[-0.340; 1.041]

Notes. Each line displays estimates from four separate OLS regressions post-processing different proxies of conditional average treatment effects (CATEs), following Chernozhukov et al. (2018). The proxies are computed using four different machine learning methods: causal forests, random forests, elastic nets and boosting. The predictors used by these algorithms are the starred variables from Table 2. The first coefficient (β_1) is an estimate of the average treatment effect of winning the lottery giving access to the FT program. The second coefficient (β_2) is an estimate of the covariate of the true CATE with its machine learning proxy (divided by the variance of the proxy). If it is not statistically different from 0, one cannot reject that the program impact is homogeneous or that the machine learning proxies are not able to capture the treatment effect heterogeneity. 90% confidence intervals are reported in square brackets following the variational inference method in Chernozhukov et al. (2018).

Source. Endline and baseline surveys.

Table 17: Impact heterogeneity: GATEs

	Causal forest			Random forest		
	γ_1	γ_5	$\gamma_1 - \gamma_5$	γ_1	γ_5	$\gamma_1 - \gamma_5$
Knowledge	0.125	0.024	-0.101	0.049	0.077	0.043
	[-0.129, 0.370]	[-0.230, 0.277]	[-0.459, 0.257]	[-0.209, 0.311]	[-0.177, 0.324]	[-0.299, 0.411]
Adoption	-0.004	0.113	0.119	0.038	0.087	0.055
	[-0.142, 0.132]	[-0.055, 0.272]	[-0.080, 0.315]	[-0.091, 0.172]	[-0.075, 0.238]	[-0.147, 0.251]
Milk production	0.079	0.258	0.154	0.145	0.163	0.008
	[-0.029, 0.194]	[-0.134, 0.685]	[-0.258, 0.609]	[-0.060, 0.364]	[-0.205, 0.529]	[-0.446, 0.449]
Profits	-0.014	0.312	0.317	0.072	0.238	0.175
	[-0.213, 0.194]	[-0.051, 0.676]	[-0.081, 0.724]	[-0.153, 0.300]	[-0.089, 0.547]	[-0.237, 0.566]
	Elastic net			Boosting		
	γ_1	γ_5	$\gamma_1 - \gamma_5$	γ_1	γ_5	$\gamma_1 - \gamma_5$
Knowledge	0.159	0.031	-0.152	0.175	-0.031	-0.193
	[-0.084, 0.402]	[-0.224, 0.288]	[-0.505, 0.207]	[-0.077, 0.424]	[-0.283, 0.212]	[-0.538, 0.157]
Adoption	-0.001	0.145	0.159	0.022	0.096	0.082
	[-0.151, 0.113]	[-0.005, 0.294]	[-0.027, 0.341]	[-0.085, 0.134]	[-0.025, 0.221]	[-0.075, 0.240]
Milk production	0.039	0.194	0.143	0.038	0.215	0.143
	[-0.099, 0.186]	[-0.126, 0.569]	[-0.247, 0.569]	[-0.018, 0.098]	[-0.195, 0.640]	[-0.426, 0.535]
Profits	-0.009	0.218	0.234	0.041	0.276	0.253
	[-0.214, 0.195]	[-0.096, 0.526]	[-0.161, 0.616]	[-0.184, 0.265]	[-0.098, 0.646]	[-0.191, 0.688]

Notes. Each line estimates of the top and bottom quintile Grouped Average Treatment Effects (GATEs), following Chernozhukov et al. (2018). Treatment effects are computed using four different machine learning methods: causal forests, random forests, elastic nets and boosting. The predictors used by these algorithms are the starred variables from Table 2. 90% confidence intervals are reported in square brackets following the variational inference method in Chernozhukov et al. (2018).

Source. Endline and baseline surveys.

Appendix

A.1 Data appendix

Data collection

The midline survey took place between July and September 2016 (approximately one year (10-16 months) after initial training), and the endline survey took place between July and September 2017 (i.e. after 2 years of program implementation). It includes questions on changes in herd composition (animal transactions, births and deaths), detailed questions on characteristics, management and production, of up to two randomly selected cows (one cross-bred and one indigenous), feed recognition with visual aids and a quiz test on knowledge of feeds and feeding practices, recent learning on feeds and practices, detailed questions on feed consumption, pro-social motivation and agency, knowledge and perceived helpfulness of the FT (asked to non-FT farmers), willingness to pay for a set of dairy production investments, knowledge and participation to the EADD program and its FT component (including trainings and visits by FT, usage and perceived efficacy of variations – this section differs for FTs and other farmers). These variables for the components to construct the knowledge, adoption, milk production and profit indices (see data appendix), and to analyze specific mechanisms.

The endline survey also included additional questions on feeds knowledge, seeds production, transfers, market supply and purchases, expectations of productivity gains from selected feeds and practices, food diversity, support received from other FTs (with the names of the specific FTs) than the one of their village, and agents/farmers involved in other EADD activities. In addition, at endline, a team of para-veterinary technicians was sent to the farmers, at the time of milking in the morning and for a subset of farmers also at time of milking in the evening, to measure directly the quantity and quality of produced milk and observe a set of indicators of animal health and weight. The visit remained non-announced to the farmer to protect against any manipulation of herd composition or milk production, which could be motivated by incorrect expectations on rewards attached to responses to the survey. Samples of produced milk (separate from both indigenous and crossbred cows) were purchased and later analyzed using a lactoscan (measurements included fat and water content of milk).[41]

[41] The survey company involved a core team of permanently-employed back checkers, supervisors and enumerators, with a college education. Some staff who worked on the baseline survey were re-employed during the midline and endline surveys, with the better performing enumerators being promoted as supervisors or back-checkers. Before each wave, field agents received an extensive two-weeks training. The PIs participated to these trainings to make sure that all questions and concepts for the survey were well-understood. Field teams were sent randomly to different areas, and in particular to treatment and control parishes. Small in-kind compensation (soap bars) was provided to farmers for participation at the end of the survey. To ensure data quality, control of durations of different parts of the interviews, tests of average values of variables corresponding to key questions (such as number of cows or plots) across enumerators, a set of consistency checks ran on a daily basis as collected data was uploaded on a server, and random short back-check interviews for a number of key questions, allocated in a way that every enumerator got checked several times, were conducted throughout

Each quantitative survey round was preceded by preparatory field work for instrument design and piloting. Three qualitative survey rounds, including focus groups and individual interviews were also conducted, taking place after the start of the intervention in the fall of 2015 and 2016 and a final round in February 2018. In addition, basic data on the number of trainers and trainees, and on the technologies discussed were reported by the FT themselves, and collected and entered through short field visits by a small separate enumerator team. Additional monitoring data is obtained from reports of the extension agents (for the linkage variation) and attendance to initial, refresher and needs assessment training, collected by the ICRAF program staff.

Indexes

Most of the variables used in the paper directly stem from the surveys,[42] and are described in the text. In addition, we construct summary indexes as weighted averages of baseline (respectively endline) variables. We set the weight of each variable in a summary index equal to the sum of its row entries in the inverted covariance matrix of all variables entering the index (see Anderson (2008), p. 1485).

At endline, we consider four main summary indexes. The knowledge index is based on 24 indicator variables equal to 1 if the farmer answered correctly the corresponding question of the knowledge test, and eight indicator variables equal to 1 if the farmer recognized the feed or feeding practice named by the enumerator (see Sections 2.7 and 2.8 of endline questionnaire for detailed questionsTo save respondents' time, (non-FT) farmers were randomly split into two samples that were only asked about a subset of the 32 feeds and feeding practices listed in the questionnaire; in the index, we only include the eight items asked to all farmers. Indexes using all available items for a given farmer make no difference. Using the exact same index for all farmers is however necessary for the analysis of treatment effect heterogeneity, as doing otherwise would mechanically create treatment effect heterogeneity.). The adoption index is based upon eight indicator variables reporting whether the eights feeds or feeding practices listed in the knowledge test were used in the past 12 months, and five indicator variables equal to 1 if the farmer reported having planted the corresponding feed on his plots over the previous 12 months (see Sections 2.8 and 3.2 of endline questionnaire). The milk production index uses two continuous variable reporting the quantity of milk produced by crossbred cows over the past three days, and over the past 12 months (see Section 2.2 of endline questionnaire). The computation of the profit index is detailed in Table 6. We select the conservative measure in which net livestock investments are subtracted from annual dairy revenues, which amounts to amortizing livestock investment over a year.

Each of the four endline summary indexes is then standardized by subtracting its mean

field work. For a small number of surveys with obvious errors or missing information, enumerators were sent back for re-collecting information.

[42]See survey instruments at http://www.parisschoolofeconomics.com/behaghel-luc/FTmaterials.zip.

value in the control group (FTs and other farmers) and dividing by its standard deviation in the control group, so that impact estimates read as standardized effects.

The construction of baseline summary indexes follows the same procedure. However, we did not gather sufficient information to compute profit measures, and the exact items entering the three remaining baseline indexes slightly differ from the endline ones, as we improved survey instruments survey after survey.[43] The baseline knowledge index is based upon the recognition of 14 feeding plants, 10 feeding/cattle management practices, and 2 commercial feeds or supplements (there was no multiple-choice question knowledge test at baseline). The baseline adoption index is based upon the reported adoption of the same 14 feeding plants, and 10 feeding/cattle management practices. Last, the milk production index is based on milk quantities reported over the past three days, and during the previous four agricultural seasons covering the past 12 months (Busoga has two wet and two dry seasons per year).

Baseline summary indexes are then standardized. As there is no distinction between treated and control farmers at baseline, we subtract the sample mean and divide by the sample standard deviation.

Cost data

The cost of the core FT program is based on two trainings of two days every year and visits of demonstration plots (on the 2nd day of trainings), and is estimated at about 110 USD (of early 2018) per FT and year. In practice every ICRAF training module was repeated with groups of 35-40 FTs trained. The cost includes: salaries, accommodation and indemnities of two ICRAF trainers, rental of a training room, small mobilization costs, payment of a transport refund, accommodation and food allowance to FTs, distribution of small quantities of seeds and seedlings (2kg of calliandra, 1 kg of lab lab, 1 bag of bracheria, 2 bags of napier grass splits per group), t-shirts and training certificates for FTs, hire of a vehicle to visit demonstration plots, and distribution of demonstration material. FTs' travel costs (transport refund, accommodation and food allowance) represent 60% of the cost. The cost of the linkage variation is based on 5 visits of a professional extension agent to each FT (which corresponds to the average number of visits received by FTs in this variation during the study), and is estimated at about 25 USD per FT and year. In practice, 15 extension agents were contracted for these visits, and each visited 14 FTs during the months the variation was in operation. Note that hiring those extension agents full time and asking them to conduct 4 visits a day would provide them with roughly the average regular salary they earn. The main cost is the payment of the visit (about 3-4 dollars), but we also include the costs of printing and distribution of vouchers and mobilization of extension agents. The cost of the needs assessment variation is based on two one-day trainings every year, and is estimated at about 47 USD per FT and year. FTs' travel costs (transport refund, accommodation and food allowance) represent 80% of this

[43] In regression involving endline profits, we use the baseline milk production index as a baseline control.

cost, with the remaining including the costs of trainers, and some small training organization costs. The cost of the sign-post variation was of about 25 USD by FT and year (50 USD for the two years of the study), and including the purchase and printing of the sign-post and its distribution together with marker pencils and a bit of cement for putting them up. In total, given that only half of FTs were allocated to each variation, the average yearly cost of a FT during the two years of the study was of about 160 USD (the cost of a FT who would receive the core program and all three variations is about 205 USD).

A.2 Cost-benefit calculation

Given uncertainty in the measurement of profits, cost-benefit analysis must be taken with caution; yet undertaking such analysis is important as it has become a key judgement criterion in the policy debate on the contact farmer extension approach. The FT program was designed as a low cost strategy to reach out to a large number of farmers. As detailed in Appendix Section A.1, the average cost of the program amounts to 160 USD per village (or FT) per year. 411 villages were covered at that average cost, depending on variations they benefited from. The costs of the linkage and signpost variations are about 25 USD per village and per year, while the cost of the needs assessment is about 50 USD per village and per year. Taking the most conservative estimates of Table 6 (in which livestock investments must be amortized over a year), we find an average impact on FTs' profits of -38,330 USHS per year, and of 81760 USHS among farmers in their village. As noted above, we do not find evidence that the treatment has a stronger effect on farmers within their network, so that aggregate impact at the village level can be obtained by adding impact on the FT and the average impact on farmers, multiplied by the average number of dairy farmers in villages, which is equal to 29. This yields an aggregate increase in dairy farmers' profits of about 1140 USD per village (=(88,970+29x146,820)/3,800). Subtracting the program costs yield a net gain of 454 USD (=614-160) per village and per year. These computations are only indicative, as the statistical precision of profit estimates is very limited. It is interesting to make a separate cost-benefit analysis for the linkage and signpost variations, as their impact on farmers' profits is statistically different from 0, and they account for a small part of the overall program costs. Based on Appendix Tables A5-A6, a similar computation implies that the estimated net benefit of the linkage variation is about 1800 USD per village and per year, and about 1668 USD for the signpost variation. Importantly, we can compute a lower bound to the net benefit of the linkage variation, by taking as conservative estimates of profit impacts the lower bound of their 90% confidence interval. The resulting lower bound is about 200 USD per year per village.

A.3 Appendix: Details on Program implementation

Attendance at the initial training of the 411 treatment FTs was complete (3 FTs had to be replaced at this stage). Attendance at the refresher trainings was also almost complete. Hence compliance was high for the core treatment. However, monitoring data reported in Table [X] shows that a share of FTs ceased to train after a few months: the rate of treatment FTs conducting at least one training was 58 percent in the period Sept-Nov 2015, 79 percent in Dec-May 2016, 62 percent in June-Nov 2016, and 60 percent in Dec-May 2017. On the other hand, active treatment FTs held multiple trainings and the intensity of their training activity remained more or less constant over time: 4 trainings, on average, over each 6 months period. The mean numbers of trainees attending those training also remained constant at about 6 until late 2017. These numbers imply more than 20,000 registered attendances of trainings.[44] [45]

Regarding variations, participation was almost complete for the linkage variation with 97 percent of linkage FTs confirming having received vouchers in June 2016 and 98 percent of them having recorded some visits by extension agents. Those FTs received in average 10.9 visits over the study period. 98 percent of FTs selected for the sign-post variation reported in December 2016 having received a signpost and 96 percent of them having put it up. But not all FTs reported their numbers of trainings and trainees on the signposts: based on pictures of sign-posts at the Dec 2016 follow-up survey, 37% of FTs with a sign-post displayed the information on numbers of trainings and trainees and another 7% had marks of previously erased information. Similarly, 82 percent of the FTs in the needs assessment variation reported in June 2016 having received the associated training, but only about half of those FTs were using in early 2016 the work plan forms they were supposed to fill to set and monitor individual technology adoption targets with their trainees, and only about 40 percent in June 2017. FTs in the signpost and needs assessment variations hence did not always make use of the tools they were provided with. The qualitative field work suggests these aspects were perceived as costly by some FTs and not completely understood by others. FTs in the linkage variation were more likely to train during the entire study (excepted in Sept - Nov 15, when vouchers were not yet widely distributed); for instance 73% of these FTs trained in the period Dec 16 - June 17 but only 55% of all FTs. These linkage FTs also trained in more diversified and more advanced technologies. For example, over the period Dec 16 - Jun 17, 62% of FTs in the linkage variation trained for cattle management, against 57% with sign-posts and 50% of FTs with needs assessment. There is no clear indication that the needs assessment was associated with a wider spectrum of technologies taught, confirming that this variation seems

[44] The FTs who remained inactive or ceased to train (for various reasons) early on were replaced by new FTs from the same villages. The replacement FTs attended refresher trainings and refresher needs assessment trainings. However, in villages that had been selected for the linkage and sign-post variations and when the replacement occurred late, the new FTs were not given sign-posts or vouchers for visits of extension agents.

[45] Some elements of reporting were not fully implemented, such as signing by trainees of attendance forms.

less effective.

Table A1: Balancing

	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N
	Farmer trainers					Other farmers				
A. Household head background										
Age	0.01	(0.85)	0.99	0.99	623	0.48	(0.5)	0.33	0.80	2478
Male	-0.02	(0.02)	0.37	0.88	623	0.00	(0.01)	0.84	0.92	2480
Education index	0.03	(0.04)	0.34	0.88	623	0.02	(0.04)	0.71	0.92	2480
Illness	-0.19	(0.3)	0.52	0.93	623	0.20	(0.21)	0.34	0.78	2480
Disability	-0.01	(0.01)	0.58	0.93	623	0.01	(0.01)	0.24	0.77	2480
B. Milk										
Milk	0.22	(0.15)	0.16	0.16	584	0.15**	(0.06)	0.02	0.04	2406
Technology use	-0.02	(0.11)	0.84	0.97	626	-0.05	(0.04)	0.26	0.66	2495
# commercial feeds	0.04	(0.08)	0.66	0.97	626	-0.03	(0.03)	0.35	0.83	2495
Technology know.	-0.04	(0.07)	0.61	0.97	626	-0.03	(0.05)	0.59	0.83	2495
Experience crossbred	0.08	(0.34)	0.82	0.97	627	0.19	(0.21)	0.36	0.83	2495
Experience indig.	1.25	(0.85)	0.15	0.73	626	-0.42	(0.45)	0.35	0.83	2494
C. Cattle										
# cross. cows	-0.04	(0.18)	0.83	0.96	627	0.09	(0.07)	0.17	0.37	2495
# indig. cows	0.09	(0.28)	0.74	0.96	627	-0.37	(0.22)	0.10	0.30	2495

Table A1: (continued)

	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N
	Farmer trainers					Other farmers				
# cross. cattle	0.04	(0.48)	0.93	0.96	627	0.25	(0.17)	0.13	0.33	2495
# indig. cattle	0.41	(0.72)	0.57	0.92	627	-0.92*	(0.51)	0.07	0.23	2495
No pasture	-0.02	(0.04)	0.50	0.92	627	-0.03	(0.02)	0.14	0.37	2495
Vaccinated	0.06**	(0.02)	0.01	0.25	627	0.00	(0.02)	0.81	0.81	2495
AI	0.06	(0.04)	0.18	0.58	627	0.02*	(0.01)	0.05	0.30	2495
D. Pasture										
Pasture land	0.91**	(0.33)	0.01	0.10	626	0.35	(0.32)	0.28	0.64	2485
Water source	0.08*	(0.04)	0.07	0.10	626	0.02	(0.03)	0.41	0.64	2495
E. Capital										
# assets	0.41**	(0.16)	0.01	0.06	627	0.01	(0.09)	0.92	0.92	2495
Land size (in acres)	1.38	(1)	0.17	0.17	623	0.94	(0.7)	0.19	0.57	2479
Cultivation land	0.76*	(0.41)	0.07	0.11	624	0.25	(0.18)	0.17	0.64	2482
F. Labor										
Hours worked livestock	0.40	(0.95)	0.67	0.71	627	0.14	(0.49)	0.77	0.79	2495

Table A2: Main results: no control for baseline values

	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N	β	Std. Error	Unadjusted p-value	Romano-Wolf p-value	N
	Farmer trainers					Other farmers				
Outcome (standardized index)										
Knowledge	0.49***	(0.06)	0.00	0.00	607	0.07*	(0.04)	0.08	0.24	2459
Adoption	0.19***	(0.04)	0.00	0.00	607	0.06**	(0.03)	0.03	0.08	2460
Milk production	0.25	(0.15)	0.10	0.19	606	0.15**	(0.06)	0.02	0.07	2459
Profits	0.15	(0.1)	0.13	0.19	607	0.09	(0.06)	0.13	0.24	2460

Table A3: FT program impact on observed cow size, health, and milk quality

	Farmer trainers			Other farmers		
	N	T	(s.e)	N	T	(s.e)
	Panel A: Impact of assignment to needs assessment variation					
Cow size and health						
average cow size (circumference in cm)	270	-1.739	(1.60)	326	-0.984	(1.30)
average index of bad health (first factor)	270	-0.102	(0.10)	326	-0.127	(0.09)
Milk quality						
fat content	324	-0.289	(0.27)	257	0.154	(0.30)
standardized fat content (accounting for breed and session)	324	-0.141	(0.13)	257	0.082	(0.14)
standardized fat content (accounting for breed and session) - excluding quality flags	303	-0.079	(0.14)	238	0.107	(0.16)
	Panel B: Impact of assignment to linkage variation					
Cow size and health						
average cow size (circumference in cm)	270	2.106*	(1.24)	326	1.216	(1.09)
average index of bad health (first factor)	270	0.067	(0.11)	326	0.094	(0.11)
Milk quality						
fat content	324	-0.246	(0.25)	257	-0.094	(0.23)
standardized fat content (accounting for breed and session)	324	-0.123	(0.11)	257	-0.043	(0.11)
standardized fat content (accounting for breed and session) - excluding quality flags	303	-0.138	(0.14)	238	-0.006	(0.12)
	Panel C: Impact of assignment to sign post					
Cow size and health						
average cow size (circumference in cm)	270	2.868**	(1.37)	326	1.626	(1.05)
average index of bad health (first factor)	270	0.048	(0.11)	326	0.089	(0.09)
Milk quality						
fat content	324	0.531**	(0.25)	257	-0.324	(0.28)
standardized fat content (accounting for breed and session)	324	0.262**	(0.12)	257	-0.178	(0.14)
standardized fat content (accounting for breed and session) - excluding quality flags	303	0.218*	(0.11)	238	-0.140	(0.14)

Table A4: Impact on profits: needs assessment variation

	Mean Control	T	Std. Error	p-value	*N*	Mean Control	T	Std. Error	p-value	*N*
	FTs					Farmers				
Detail on annual revenues and annual expenditures										
Annual revenues from milk production and other earnings	2715.92	-97.25	(229.44)	0.67	395	1967.57	84.86	(130.27)	0.52	1596
Milk production	1856.39	-230.06	(156.93)	0.15	395	1232.13	56.35	(92.15)	0.54	1596
Livestock sales	766.17	148.70	(139.68)	0.29	395	647.17	72.34	(63.57)	0.26	1596
Earnings from manure/oxen/fees for impregnation	93.36	-15.89	(27.03)	0.56	395	88.27	-43.83	(28.39)	0.13	1596
Annual expenditures	826.45	55.32	(107.82)	0.61	395	493.18	34.88	(50.21)	0.49	1596
Expenditures: Feeds/feeding practices	201.88	1.69	(49.71)	0.97	395	58.24	2.44	(14.83)	0.87	1596
Expenditures: Seeds	3.96	-2.98***	(1.01)	0.00	395	0.82	0.36	(0.65)	0.58	1596
Expenditures: Health	267.08	-41.08*	(21.66)	0.06	395	180.24	4.00	(12.57)	0.75	1596
Expenditures: Livestock acquisition	175.94	107.08	(82.61)	0.20	395	150.80	18.97	(26.84)	0.48	1596
Expenditures: Hired labor	110.58	-0.03	(23.72)	1.00	395	68.81	-6.53	(10.58)	0.54	1596
Expenditures: Fencing/paddocking/barns/storage	33.39	-1.98	(12.21)	0.87	395	11.60	12.74	(10.81)	0.24	1596
Expenditures: Extension advice	0.71	-0.19	(0.46)	0.68	395	0.16	0.08	(0.1)	0.42	1596
Expenditures: Compensation for animal destruction	36.88	-10.16	(7.71)	0.19	395	23.32	3.17	(4.16)	0.45	1596
Annual profits										
Excluding livestock sales/acquisition and feed expenses	1501.11	-192.50	(148.14)	0.2	395	1036.26	-0.95	(86.65)	0.99	1596
With livestock sales/acquisition but excluding feed expenses	1299.24	-194.19	(158.96)	0.23	395	978.02	-3.39	(84.09)	0.97	1596
With livestock sales/acquisition and feed expenses	1889.47	-152.57	(249.49)	0.54	395	1474.39	49.98	(121.12)	0.68	1596

Table A5: Impact on profits: linkage variation

	Mean Control	T	Std. Error	p-value	*N*	Mean Control	T	Std. Error	p-value	*N*
	FTs					Farmers				
Detail on annual revenues and annual expenditures										
Annual revenues from milk production and other earnings	2585.90	170.75	(233.83)	0.47	395	1919.94	236.33*	(125.85)	0.07	1596
Milk production	1659.44	154.25	(199.95)	0.44	395	1161.02	217.83**	(85.98)	0.01	1596
Livestock sales	838.76	20.15	(96.2)	0.83	395	701.62	-1.85	(54.68)	0.97	1596
Earnings from manure/oxen/fees for impregnation	87.70	-3.65	(26.06)	0.89	395	57.31	20.35	(21.19)	0.34	1596
Annual expenditures	890.06	-82.04	(130.56)	0.53	395	500.88	28.85	(54.5)	0.60	1596
Expenditures: Feeds/feeding practices	227.15	-67.96	(44.91)	0.14	395	52.67	12.26	(15.13)	0.42	1596
Expenditures: Seeds	3.05	-0.99	(1.14)	0.39	395	0.98	-0.01	(0.63)	0.99	1596
Expenditures: Health	228.44	32.28	(31.42)	0.31	395	168.43	30.06**	(11.5)	0.01	1596
Expenditures: Livestock acquisition	241.94	-15.53	(76.56)	0.84	395	169.55	-12.23	(24.62)	0.62	1596
Expenditures: Hired labor	107.18	11.03	(29.11)	0.71	395	68.33	-4.47	(11.69)	0.70	1596
Expenditures: Fencing/paddocking/barns/storage	48.46	-30.66**	(12.13)	0.01	395	16.71	2.17	(12.32)	0.86	1596
Expenditures: Extension advice	0.99	-0.81	(0.62)	0.20	395	0.18	0.04	(0.12)	0.74	1596
Expenditures: Compensation for animal destruction	35.90	-10.40	(10.55)	0.33	395	25.01	1.01	(3.33)	0.76	1596
Annual profits										
Excluding livestock sales/acquisition and feed expenses	1326.16	149.15	(176.3)	0.4	395	939.66	209.36**	(75.77)	0.01	1596
With livestock sales/acquisition but excluding feed expenses	1099.01	217.12	(176.78)	0.22	395	887	197.10**	(73.35)	0.01	1596
With livestock sales/acquisition and feed expenses	1695.83	252.79	(231.41)	0.28	395	1419.06	207.48*	(104.34)	0.05	1596

Table A6: Impact on profits: signposts variation

	Mean Control	T	Std. Error	p-value	*N*	Mean Control	T	Std. Error	p-value	*N*
	FTs					Farmers				
Detail on annual revenues and annual expenditures										
Annual revenues from milk production and other earnings	2577.60	189.81	(277.5)	0.50	395	1930.96	307.63**	(117.2)	0.01	1596
Milk production	1634.05	199.20	(200.39)	0.32	395	1210.85	199.83**	(91.19)	0.03	1596
Livestock sales	863.56	-24.33	(126.79)	0.85	395	667.87	74.03	(50.75)	0.15	1596
Earnings from manure/oxen/fees for impregnation	79.99	14.93	(26.78)	0.58	395	52.24	33.78	(25.59)	0.19	1596
Annual expenditures	832.09	-5.70	(142.77)	0.97	395	472.60	84.86**	(36.53)	0.02	1596
Expenditures: Feeds/feeding practices	219.49	-59.58	(57.81)	0.31	395	51.74	19.15	(11.78)	0.11	1596
Expenditures: Seeds	1.78	1.41	(1.29)	0.28	395	0.47	1.07	(0.64)	0.10	1596
Expenditures: Health	235.82	14.23	(27.01)	0.60	395	181.62	9.05	(10.07)	0.37	1596
Expenditures: Livestock acquisition	208.88	39.12	(76.6)	0.61	395	144.27	22.79	(24.58)	0.36	1596
Expenditures: Hired labor	105.57	4.96	(29.39)	0.87	395	61.49	12.89	(9.97)	0.20	1596
Expenditures: Fencing/paddocking/barns/storage	32.87	-11.24	(12.13)	0.36	395	8.62	19.87*	(9.73)	0.05	1596
Expenditures: Extension advice	0.79	-0.18	(0.53)	0.74	395	0.11	0.20*	(0.11)	0.07	1596
Expenditures: Compensation for animal destruction	28.68	6.98	(8.24)	0.40	395	24.74	0.91	(3.03)	0.77	1596
Annual profits										
Excluding livestock sales/acquisition and feed expenses	1310.32	199.38	(186.57)	0.29	395	986.5	190.69**	(86.21)	0.03	1596
With livestock sales/acquisition but excluding feed expenses	1090.83	258.95	(183.94)	0.17	395	934.76	171.54*	(85.33)	0.05	1596
With livestock sales/acquisition and feed expenses	1745.51	195.51	(268.45)	0.47	395	1458.36	222.77*	(111.46)	0.05	1596

Table A7: FT program impact on willingness to pay for animal, investments in animals, and feeds and practices

	Mean		Std.	Unadj.	R-Wolf		Mean		Std.	Unadj.	R-Wolf	
	control	β	Error	p-value	p-value	*N*	control	β	Error	p-value	p-value	*N*
			Farmer trainers						Other farmers			
Outcome (standardized indexes)							A. Main treatment					
Willingness to pay for cows and heifers	-0.02	0.02	(0.09)	0.78	0.77	607	0.01	-0.05	(0.04)	0.17	0.42	2460
Willingness to pay for animal health and AI	-0.05	-0.20**	(0.07)	0.01	0.04	607	0.01	-0.02	(0.05)	0.72	0.76	2460
Willingness to pay for feeds and feeding practices	0.18	0.18**	(0.08)	0.03	0.10	607	-0.04	0.08*	(0.04)	0.09	0.23	2460
							B. Needs assessment					
Willingness to pay for cows and heifers	-0.02	0.10	(0.1)	0.30	0.67	395	-0.09	0.10**	(0.05)	0.04	0.09	1596
Willingness to pay for animal health and AI	-0.27	0.03	(0.1)	0.79	0.79	395	-0.02	0.02	(0.05)	0.72	0.91	1596
Willingness to pay for feeds and feeding practices	0.31	0.10	(0.12)	0.37	0.67	395	0.01	0.01	(0.05)	0.89	0.91	1596
							C. Linkage					
Willingness to pay for cows and heifers	-0.02	0.08	(0.1)	0.39	0.64	395	-0.02	-0.04	(0.05)	0.41	0.66	1596
Willingness to pay for animal health and AI	-0.27	0.03	(0.1)	0.76	0.75	395	0	-0.01	(0.05)	0.78	0.78	1596
Willingness to pay for feeds and feeding practices	0.46	-0.20*	(0.12)	0.09	0.26	395	-0.04	0.09*	(0.05)	0.06	0.18	1596
							D. Signpost					
Willingness to pay for cows and heifers	0.02	0.01	(0.1)	0.93	0.98	395	-0.02	-0.03	(0.05)	0.58	0.82	1596
Willingness to pay for animal health and AI	-0.24	-0.02	(0.1)	0.84	0.98	395	-0.01	0.01	(0.05)	0.89	0.90	1596
Willingness to pay for feeds and feeding practices	0.3	0.11	(0.12)	0.32	0.69	395	0.05	-0.07	(0.05)	0.13	0.34	1596

Table A8: FT program impact on general interactions

	Mean		Std.	Unadj.	R-Wolf		Mean		Std.	Unadj.	R-Wolf	
	control	β	Error	p-value	p-value	N	control	β	Error	p-value	p-value	N
			Farmer trainers						Other farmers			
Outcome							A. Main treatment					
Household holds a local leadership position	0.9	0.04	(0.03)	0.21	0.35	607	0.39	0.00	(0.02)	0.85	0.86	2460
Household participates in credit/saving group	0.67	0.00	(0.04)	0.94	0.94	607	0.52	0.02	(0.02)	0.45	0.62	2460
							B. Needs assessment					
Household holds a local leadership position	0.93	0.03	(0.02)	0.29	0.50	395	0.39	-0.01	(0.02)	0.83	0.97	1596
Household participates in credit/saving group	0.68	-0.04	(0.05)	0.45	0.50	395	0.54	0.00	(0.02)	0.96	0.97	1596
							C. Linkage					
Household holds a local leadership position	0.92	0.05*	(0.02)	0.05	0.10	395	0.38	0.03	(0.02)	0.23	0.42	1596
Household participates in credit/saving group	0.66	-0.01	(0.05)	0.81	0.80	395	0.54	-0.02	(0.02)	0.52	0.53	1596
							D. Signpost					
Household holds a local leadership position	0.93	0.02	(0.02)	0.45	0.69	395	0.41	-0.04	(0.02)	0.13	0.11	1596
Household participates in credit/saving group	0.64	0.03	(0.05)	0.52	0.69	395	0.51	0.05*	(0.02)	0.06	0.11	1596

Table A9: FT program impact on feeds purchases

	Mean		Std.	Unadj.	R-Wolf		Mean		Std.	Unadj.	R-Wolf	
	control	β	Error	p-value	p-value	N	control	β	Error	p-value	p-value	N
			Farmer trainers						Other farmers			
Outcome							A. Main treatment					
Purchased any feed (past 3 months)	0.54	0.08*	(0.04)	0.05	0.11	607	0.4	0.01	(0.02)	0.78	0.85	2460
Nb of feeds purchased (past 3 months)	1.13	0.74***	(0.14)	0.00	0.01	607	0.66	0.05	(0.05)	0.35	0.74	2460
Total cost of feeds purchased (past 3 months, th. Ushs)	43.8	47.75**	(16.86)	0.01	0.11	607	27.66	3.52	(4.58)	0.44	0.85	2460
Purchased any feed (last dry season)	0.56	0.02	(0.04)	0.55	0.70	607	0.37	0.01	(0.02)	0.68	0.85	2460
Nb of feeds purchased (last dry season)	1.14	0.61***	(0.16)	0.00	0.01	607	0.63	0.06	(0.05)	0.27	0.68	2460
Total cost of feeds purchased (last dry season, th. Ushs)	72.96	15.51	(19.02)	0.42	0.70	607	26.28	8.20	(5.35)	0.13	0.65	2460
							B. Needs assessment					
Purchased any feed (past 3 months)	0.67	-0.09*	(0.05)	0.07	0.28	395	0.41	-0.01	(0.02)	0.77	0.93	1596
Nb of feeds purchased (past 3 months)	2.06	-0.26	(0.24)	0.27	0.69	395	0.73	-0.07	(0.06)	0.25	0.64	1596
Total cost of feeds purchased (past 3 months, th. Ushs)	100.94	-14.67	(31.8)	0.64	0.93	395	29.12	2.31	(6.72)	0.73	0.93	1596
Purchased any feed (last dry season)	0.59	-0.02	(0.05)	0.72	0.93	395	0.39	-0.01	(0.02)	0.56	0.93	1596
Nb of feeds purchased (last dry season)	1.86	-0.18	(0.23)	0.44	0.85	395	0.69	-0.04	(0.06)	0.55	0.93	1596
Total cost of feeds purchased (last dry season, th. Ushs)	95.05	-11.96	(31.77)	0.71	0.93	395	26.92	14.40	(8.94)	0.11	0.44	1596
							C. Linkage					
Purchased any feed (past 3 months)	0.61	0.02	(0.05)	0.66	0.90	395	0.39	0.03	(0.02)	0.25	0.62	1596
Nb of feeds purchased (past 3 months)	1.8	0.22	(0.24)	0.35	0.69	395	0.67	0.05	(0.06)	0.45	0.70	1596
Total cost of feeds purchased (past 3 months, th. Ushs)	113.57	-39.00	(31.74)	0.22	0.57	395	26.33	7.32	(6.72)	0.28	0.62	1596
Purchased any feed (last dry season)	0.53	0.11**	(0.05)	0.03	0.18	395	0.36	0.03	(0.02)	0.17	0.53	1596
Nb of feeds purchased (last dry season)	1.44	0.65**	(0.23)	0.01	0.07	395	0.62	0.10	(0.06)	0.14	0.48	1596
Total cost of feeds purchased (last dry season, th. Ushs)	88.59	2.04	(31.77)	0.95	0.96	395	32.1	3.15	(8.95)	0.72	0.73	1596
							D. Signpost					
Purchased any feed (past 3 months)	0.6	0.05	(0.05)	0.30	0.71	395	0.4	0.01	(0.02)	0.66	0.65	1596
Nb of feeds purchased (past 3 months)	1.89	0.05	(0.24)	0.83	0.97	395	0.65	0.08	(0.06)	0.16	0.49	1596
Total cost of feeds purchased (past 3 months, th. Ushs)	109.75	-32.57	(31.77)	0.31	0.71	395	25.87	8.96	(6.72)	0.18	0.49	1596
Purchased any feed (last dry season)	0.58	0.00	(0.05)	0.96	0.97	395	0.36	0.04*	(0.02)	0.07	0.33	1596
Nb of feeds purchased (last dry season)	1.95	-0.34	(0.23)	0.15	0.53	395	0.62	0.11	(0.06)	0.10	0.37	1596
Total cost of feeds purchased (last dry season, th. Ushs)	107.42	-37.67	(31.72)	0.24	0.67	395	30.18	7.61	(8.95)	0.40	0.64	1596

Table A10: FT program impact on investments in animals

	Mean		Std.	Unadj.	R-Wolf		Mean		Std.	Unadj.	R-Wolf	
	control	β	Error	p-value	p-value	N	control	β	Error	p-value	p-value	N
			Farmer trainers						Other farmers			
Outcome							A. Main treatment					
Average price of exotic/crossbreed cows (th. Ushs)	1325.6	46.53	(80.62)	0.57	0.82	151	1178.69	66.67	(54.17)	0.22	0.79	396
Average price of indigenous cows (th. Ushs)	906.27	-33.26	(34.26)	0.34	0.82	180	828.33	25.58	(18.47)	0.17	0.68	801
Expanses in AI for exotic cows (last 12 months, th. Ushs)	1.01	1.30	(0.95)	0.18	0.79	255	0.67	0.23	(0.47)	0.63	0.95	710
Expanses in AI for indigenous cows (last 12 months, th. Ushs)	0.09	-0.09	(0.08)	0.25	0.74	296	0.17	-0.03	(0.12)	0.79	0.96	1391
Hours spent on usual milk production activities (last week)	145.8	-39.61	(52.15)	0.45	0.82	607	76.72	-5.42	(5.31)	0.31	0.79	2460
Hours spent on clearing the land for pasturing/fencing paddocks (last 12 months)	3.64	2.37	(1.9)	0.22	0.81	607	6.26	0.65	(2.3)	0.78	0.96	2460
							B. Needs assessment					
Average price of exotic/crossbreed cows (th. Ushs)	1363.9	-79.46	(99.22)	0.43	0.67	96	1250.47	-2.57	(75.37)	0.97	1.00	258
Average price of indigenous cows (th. Ushs)	857.68	38.80	(58.51)	0.51	0.67	121	844.9	9.35	(21.23)	0.66	0.98	515
Expanses in AI for exotic cows (last 12 months, th. Ushs)	3.47	-2.16	(1.47)	0.14	0.50	171	0.58	0.45	(0.53)	0.39	0.95	475
Expanses in AI for indigenous cows (last 12 months, th. Ushs)	0	0.00	(0)	.	0.00	188	0.13	-0.01	(0.14)	0.96	1.00	892
Hours spent on usual milk production activities (last week)	88.81	21.78	(14.38)	0.13	0.50	395	73.69	-2.47	(4.71)	0.60	0.98	1596
Hours spent on clearing the land for pasturing/fencing paddocks (last 12 months)	4.05	5.61	(3.4)	0.10	0.50	395	5.75	2.47	(3.15)	0.43	0.96	1596
							C. Linkage					
Average price of exotic/crossbreed cows (th. Ushs)	1360.6	-49.52	(98.85)	0.62	0.91	96	1229.63	20.45	(76.44)	0.79	0.98	258
Average price of indigenous cows (th. Ushs)	886.07	-15.97	(57.2)	0.78	0.91	121	830.64	38.88*	(21.29)	0.07	0.46	515
Expanses in AI for exotic cows (last 12 months, th. Ushs)	2.8	-0.96	(1.48)	0.52	0.91	171	0.79	0.13	(0.53)	0.81	0.98	475
Expanses in AI for indigenous cows (last 12 months, th. Ushs)	0	0.00	(0)	.	0.00	188	0.15	-0.05	(0.14)	0.74	0.98	892
Hours spent on usual milk production activities (last week)	93.26	13.97	(14.4)	0.33	0.89	395	69.66	5.45	(4.71)	0.25	0.68	1596
Hours spent on clearing the land for pasturing/fencing paddocks (last 12 months)	8.2	-3.01	(3.41)	0.38	0.89	395	8.97	-4.09	(3.15)	0.19	0.68	1596
							D. Signpost					
Average price of exotic/crossbreed cows (th. Ushs)	1434.5	-176.36*	(95.57)	0.07	0.35	96	1209.66	33.14	(75.27)	0.66	0.94	258
Average price of indigenous cows (th. Ushs)	841.47	44.81	(56.97)	0.43	0.75	121	846.96	3.87	(21.08)	0.85	0.94	515
Expanses in AI for exotic cows (last 12 months, th. Ushs)	2.77	-0.49	(1.48)	0.74	0.75	171	1.11	-0.55	(0.53)	0.30	0.85	475
Expanses in AI for indigenous cows (last 12 months, th. Ushs)	0	0.00	(0)	.	0.00	188	0.16	-0.07	(0.14)	0.60	0.94	892
Hours spent on usual milk production activities (last week)	92.49	15.55	(14.4)	0.28	0.75	395	66.03	12.41**	(4.7)	0.01	0.17	1596
Hours spent on clearing the land for pasturing/fencing paddocks (last 12 months)	8.64	-3.62	(3.41)	0.29	0.75	395	5.55	2.72	(3.15)	0.39	0.87	1596

Table A11: FT program impact on transactions of animals

	Mean		Std.	Unadj.	R-Wolf		Mean		Std.	Unadj.	R-Wolf	
	control	β	Error	p-value	p-value	*N*	control	β	Error	p-value	p-value	*N*
			Farmer trainers						Other farmers			
Outcome							A. Main treatment					
Obtained any new crossbred livestock (last year)	0.35	0.05	(0.04)	0.16	0.65	607	0.28	0.03*	(0.02)	0.09	0.36	2460
Obtained any new indigenous livestock (last year)	0.37	0.02	(0.04)	0.59	0.93	607	0.36	0.02	(0.02)	0.33	0.69	2460
Gave/sold/exchanged away some crossbred animals (last year)	0.25	0.02	(0.03)	0.41	0.93	607	0.18	0.03	(0.02)	0.11	0.33	2460
Gave/sold/exchanged away some indigenous animals (last year)	0.43	-0.02	(0.03)	0.48	0.93	607	0.42	-0.02	(0.02)	0.40	0.69	2460
Net value of livestock transactions (last year)	493.44	-17.12	(132.61)	0.90	0.93	607	265.97	55.44	(64.9)	0.40	0.69	2460
Value of livestock bought (last year)	144928	95 167.68*	(48023.26)	0.05	0.53	607	112336	49 435.92**	(17756.4)	0.01	0.10	2460
							B. Needs assessment					
Obtained any new crossbred livestock (last year)	0.42	-0.02	(0.05)	0.69	0.97	395	0.32	-0.02	(0.02)	0.30	0.68	1596
Obtained any new indigenous livestock (last year)	0.39	0.01	(0.05)	0.89	0.97	395	0.39	-0.03	(0.02)	0.23	0.68	1596
Gave/sold/exchanged away some crossbred animals (last year)	0.23	0.08*	(0.04)	0.06	0.33	395	0.19	0.03	(0.02)	0.15	0.61	1596
Gave/sold/exchanged away some indigenous animals (last year)	0.4	-0.01	(0.05)	0.80	0.97	395	0.41	0.00	(0.02)	0.85	0.86	1596
Net value of livestock transactions (last year)	398.46	204.98	(159.7)	0.20	0.70	395	261.4	92.35	(72.2)	0.20	0.68	1596
Value of livestock bought (last year)	175939	104 983.27	(88522.05)	0.24	0.70	395	150801	15 592.87	(26944.8)	0.56	0.83	1596
							C. Linkage					
Obtained any new crossbred livestock (last year)	0.44	-0.07	(0.05)	0.16	0.66	395	0.29	0.03	(0.02)	0.19	0.69	1596
Obtained any new indigenous livestock (last year)	0.41	-0.03	(0.05)	0.58	0.97	395	0.39	-0.03	(0.02)	0.21	0.69	1596
Gave/sold/exchanged away some crossbred animals (last year)	0.27	0.01	(0.04)	0.79	1.00	395	0.19	0.01	(0.02)	0.45	0.71	1596
Gave/sold/exchanged away some indigenous animals (last year)	0.4	-0.01	(0.05)	0.85	1.00	395	0.43	-0.03	(0.02)	0.19	0.69	1596
Net value of livestock transactions (last year)	418.98	172.01	(159.79)	0.28	0.80	395	259.76	84.92	(72.2)	0.24	0.69	1596
Value of livestock bought (last year)	241939	-25 496.36	(88667.31)	0.77	1.00	395	169552	-15 196.28	(26944.2)	0.57	0.71	1596
							D. Signpost					
Obtained any new crossbred livestock (last year)	0.35	0.11**	(0.05)	0.02	0.15	395	0.27	0.06**	(0.02)	0.01	0.03	1596
Obtained any new indigenous livestock (last year)	0.37	0.05	(0.05)	0.36	0.83	395	0.38	-0.01	(0.02)	0.68	0.97	1596
Gave/sold/exchanged away some crossbred animals (last year)	0.3	-0.04	(0.04)	0.31	0.83	395	0.19	0.02	(0.02)	0.25	0.70	1596
Gave/sold/exchanged away some indigenous animals (last year)	0.39	0.01	(0.05)	0.87	0.87	395	0.41	0.00	(0.02)	0.85	0.97	1596
Net value of livestock transactions (last year)	585.77	-154.40	(159.88)	0.33	0.83	395	315.64	-18.74	(72.23)	0.80	0.97	1596
Value of livestock bought (last year)	208878	41 390.14	(88679.18)	0.64	0.87	395	144274	31 446.14	(26935.13)	0.24	0.70	1596

Table A12: Intra-village and intra-parish correlations

	Intra-village correlation coefficient		Intra-parish correlation coefficient	
	FTs and other farmers	Other farmers only	FTs and other farmers	Other farmers only
Milk production index	0.12	0.08	0.05	0.05
	[0.08 ; 0.15]	[0.05 ; 0.12]	[0.03 ; 0.08]	[0.03 ; 0.08]
Technology usage index	0.27	0.47	0.15	0.21
	[0.23 ; 0.3]	[0.43 ; 0.51]	[0.1 ; 0.2]	[0.15 ; 0.28]
Number of commercial feeds used	0.25	0.41	0.06	0.08
	[0.21 ; 0.29]	[0.37 ; 0.46]	[0.03 ; 0.08]	[0.05 ; 0.11]
Technology knowledge index	0.26	0.39	0.11	0.12
	[0.22 ; 0.3]	[0.34 ; 0.43]	[0.07 ; 0.15]	[0.07 ; 0.16]
Years of experience with cross-bred cattle	0.17	0.17	0.11	0.10
	[0.14 ; 0.21]	[0.13 ; 0.21]	[0.07 ; 0.14]	[0.06 ; 0.14]
Years of experience with indigenous cattle	0.07	0.07	0.04	0.03
	[0.04 ; 0.1]	[0.03 ; 0.11]	[0.02 ; 0.06]	[0.01 ; 0.05]
# crossbred cows	0.09	0.12	0.05	0.08
	[0.06 ; 0.12]	[0.09 ; 0.16]	[0.03 ; 0.08]	[0.05 ; 0.11]
# indigenous cows	0.14	0.15	0.11	0.12
	[0.11 ; 0.18]	[0.11 ; 0.19]	[0.07 ; 0.15]	[0.07 ; 0.16]
# crossbred heads of cattle*	0.10	0.16	0.05	0.08
	[0.07 ; 0.13]	[0.12 ; 0.2]	[0.03 ; 0.07]	[0.05 ; 0.12]
# indigenous heads of cattle*	0.18	0.19	0.15	0.15
	[0.15 ; 0.22]	[0.15 ; 0.23]	[0.1 ; 0.2]	[0.1 ; 0.2]
# cows per acre of land	0.09	0.11	0.05	0.06
	[0.06 ; 0.12]	[0.07 ; 0.15]	[0.03 ; 0.08]	[0.03 ; 0.08]
=1 if no pasture	0.10	0.14	0.03	0.03
	[0.07 ; 0.13]	[0.1 ; 0.18]	[0.01 ; 0.05]	[0.01 ; 0.05]
=1 if cattle is vaccinated	0.05	0.05	0.03	0.03
	[0.02 ; 0.08]	[0.02 ; 0.09]	[0.01 ; 0.05]	[0.01 ; 0.05]
=1 if used artificial insemination or controlled breeding	0.06	0.15	0.03	0.05
	[0.03 ; 0.09]	[0.11 ; 0.18]	[0.01 ; 0.05]	[0.02 ; 0.07]
Pasture land size (acres)*	0.14	0.17	0.07	0.07
	[0.11 ; 0.18]	[0.12 ; 0.21]	[0.04 ; 0.1]	[0.04 ; 0.11]
=1 if water source on pasture*	0.16	0.23	0.04	0.05
	[0.13 ; 0.2]	[0.19 ; 0.27]	[0.02 ; 0.06]	[0.02 ; 0.07]
# assets held on list 24*	0.09	0.12	0.03	0.04
	[0.06 ; 0.12]	[0.09 ; 0.16]	[0.02 ; 0.05]	[0.02 ; 0.06]
Land size (in acres)*	0.12	0.13	0.06	0.07
	[0.09 ; 0.16]	[0.09 ; 0.17]	[0.03 ; 0.09]	[0.04 ; 0.1]
Cultivation land size (in acres)*	0.08	0.11	0.04	0.04
	[0.05 ; 0.11]	[0.07 ; 0.15]	[0.02 ; 0.05]	[0.01 ; 0.06]
Total hours worked with livestock in a day*	0.23	0.29	0.06	0.06
	[0.2 ; 0.27]	[0.25 ; 0.33]	[0.04 ; 0.09]	[0.03 ; 0.08]

Table A13: Intra-village correlation at endline: treatment vs. control

	Treatment villages		Control villages	
	FTs and other farmers	Other farmers only	FTs and other farmers	Other farmers only
Knowledge	0.00	0.05	0.05	0.08
	[0 ; 0.04]	[0.01 ; 0.09]	[0 ; 0.1]	[0.02 ; 0.14]
Adoption	0.14	0.21	0.15	0.17
	[0.1 ; 0.18]	[0.16 ; 0.26]	[0.09 ; 0.21]	[0.1 ; 0.24]
Milk production	0.13	0.09	0.05	0.06
	[0.09 ; 0.17]	[0.04 ; 0.14]	[0 ; 0.1]	[0 ; 0.12]
Profits	0.10	0.09	0.05	0.08
	[0.06 ; 0.14]	[0.04 ; 0.14]	[0 ; 0.1]	[0.01 ; 0.14]

Table A14: Spillover analysis: diffusion channels

	Knows katiti as feed			Has received training on dairy farming			Has been given seeds for free			Has been given seeds by other farmer		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.047***	0.042*	0.046***	0.291***	0.227***	0.295***	0.096***	0.123***	0.103***	0.023	0.022	0.041
	(0.012)	(0.024)	(0.015)	(0.027)	(0.059)	(0.033)	(0.019)	(0.035)	(0.022)	(0.047)	(0.101)	(0.071)
NT_{02}		0.001			0.009			0.003			-0.000	
		(0.011)			(0.017)			(0.011)			(0.032)	
NT_{25}		0.001			0.003			0.000			0.012	
		(0.002)			(0.005)			(0.003)			(0.009)	
$NT_{02} \times T$		-0.000			0.004			-0.005			-0.009	
		(0.010)			(0.018)			(0.012)			(0.036)	
$NT_{25} \times T$		-0.000			0.001			-0.001			0.001	
		(0.002)			(0.004)			(0.003)			(0.009)	
$T_{02} \times (1-T)$			-0.002			0.008			0.013			0.033
			(0.020)			(0.034)			(0.025)			(0.088)
Observations	2,460	2,460	2,460	2,460	2,460	2,460	2,460	2,460	2,460	418	418	418
Mean outcome in control	0.097	0.097	0.097	0.264	0.264	0.264	0.116	0.116	0.116	0.788	0.788	0.788
Implied total effect	0.031	0.040	0.029	0.190	0.206	0.194	0.063	0.061	0.070	0.015	0.128	0.032
se	0.008	0.026	0.012	0.018	0.055	0.025	0.012	0.033	0.018	0.031	0.108	0.059
p-value total effet	0.000	0.130	0.015	0.000	0.000	0.000	0.000	0.069	0.000	0.628	0.240	0.583
Implied spillover effect		0.012	0.000		0.058	0.001		-0.019	0.002		0.113	0.006
se		0.026	0.004		0.060	0.006		0.036	0.004		0.095	0.015
p-value spillover effect		0.631	0.903		0.334	0.810		0.597	0.602		0.237	0.713

Table A15: Spillover analysis among network farmers: main outcomes

	Knowledge			Adoption			Milk production			Profits		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.070	0.194*	0.046	0.068**	0.118**	0.107***	0.163**	0.285**	0.170**	0.088	0.042	0.094
	(0.048)	(0.098)	(0.064)	(0.031)	(0.055)	(0.040)	(0.064)	(0.115)	(0.074)	(0.061)	(0.113)	(0.066)
NT_{02}		-0.020			-0.018			-0.013			0.018	
		(0.048)			(0.024)			(0.037)			(0.041)	
NT_{25}		-0.001			0.011			0.012			0.008	
		(0.010)			(0.006)			(0.009)			(0.011)	
$NT_{02} \times T$		-0.006			0.018			-0.047			-0.021	
		(0.043)			(0.024)			(0.032)			(0.042)	
$NT_{25} \times T$		-0.002			-0.008			0.007			0.006	
		(0.008)			(0.005)			(0.008)			(0.009)	
$T_{02} \times (1 - T)$			-0.048			0.078			0.013			0.013
			(0.099)			(0.054)			(0.062)			(0.063)
Observations	1,852	1,852	1,852	1,853	1,853	1,853	1,852	1,852	1,852	1,853	1,853	1,853
Mean outcome in control	-0.091	-0.091	-0.091	-0.085	-0.085	-0.085	-0.040	-0.040	-0.040	0.006	0.006	0.006
Implied total effect	0.046	0.058	0.022	0.045	0.150	0.083	0.106	0.240	0.113	0.057	0.123	0.064
se	0.031	0.131	0.055	0.020	0.076	0.033	0.041	0.103	0.055	0.040	0.128	0.048
p-value total effet	0.148	0.659	0.692	0.028	0.050	0.013	0.012	0.022	0.041	0.155	0.340	0.191
Implied spillover effect		-0.068	-0.008		0.073	0.013		0.054	0.002		0.096	0.002
se		0.119	0.017		0.065	0.009		0.093	0.011		0.134	0.011
p-value spillover effect		0.569	0.629		0.264	0.152		0.562	0.831		0.477	0.840

Table A16: Spillover analysis among network farmers: diffusion channels

	Knows katiti as feed			Has received training on dairy farming			Has been given seeds for free			Has been given seeds by other farmer		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.057***	0.046*	0.038*	0.301***	0.228***	0.316***	0.102***	0.104***	0.105***	-0.020	0.055	0.040
	(0.013)	(0.025)	(0.020)	(0.029)	(0.063)	(0.034)	(0.020)	(0.039)	(0.023)	(0.056)	(0.099)	(0.092)
NT_{02}		-0.012			0.021			0.008			0.008	
		(0.012)			(0.019)			(0.014)			(0.035)	
NT_{25}		0.000			0.002			0.001			0.015	
		(0.002)			(0.005)			(0.004)			(0.009)	
$NT_{02} \times T$		0.010			-0.005			-0.003			-0.028	
		(0.012)			(0.020)			(0.015)			(0.038)	
$NT_{25} \times T$		0.001			0.002			-0.001			-0.001	
		(0.002)			(0.005)			(0.003)			(0.009)	
$T_{02} \times (1 - T)$			-0.037			0.031			0.006			0.114
			(0.028)			(0.036)			(0.028)			(0.103)
Observations	1,853	1,853	1,853	1,853	1,853	1,853	1,853	1,853	1,853	325	325	325
Mean outcome in control	0.093	0.093	0.093	0.273	0.273	0.273	0.118	0.118	0.118	0.816	0.816	0.816
Implied total effect	0.037	0.054	0.019	0.196	0.199	0.212	0.066	0.065	0.070	-0.013	0.112	0.046
se	0.009	0.027	0.017	0.019	0.062	0.026	0.013	0.040	0.018	0.036	0.102	0.075
p-value total effet	0.000	0.051	0.275	0.000	0.002	0.000	0.000	0.109	0.000	0.725	0.275	0.546
Implied spillover effect		0.024	-0.006		0.050	0.005		-0.003	0.001		0.076	0.020
se		0.029	0.005		0.066	0.006		0.042	0.005		0.102	0.018
p-value spillover effect		0.410	0.193		0.446	0.394		0.940	0.822		0.458	0.274

Table A17: Spillover analysis among non-network farmers: main outcomes

	Knowledge			Adoption			Milk production			Profits		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.088	0.129	0.194	0.024	0.092	0.036	0.097	0.187	0.095	0.101	0.351	0.097
	(0.071)	(0.163)	(0.144)	(0.035)	(0.060)	(0.049)	(0.086)	(0.191)	(0.109)	(0.100)	(0.229)	(0.131)
NT_{02}		0.108			0.017			-0.011			-0.037	
		(0.066)			(0.031)			(0.043)			(0.045)	
NT_{25}		0.021			0.011			0.018*			0.032**	
		(0.014)			(0.007)			(0.010)			(0.014)	
$NT_{02} \times T$		-0.090			-0.039			-0.014			-0.023	
		(0.064)			(0.031)			(0.041)			(0.048)	
$NT_{25} \times T$		-0.005			0.001			-0.001			-0.006	
		(0.013)			(0.005)			(0.011)			(0.015)	
$T_{02} \times (1 - T)$			0.201			0.022			-0.005			-0.006
			(0.202)			(0.077)			(0.122)			(0.126)
Observations	607	607	607	607	607	607	607	607	607	607	607	607
Mean outcome in control	-0.140	-0.140	-0.140	-0.094	-0.094	-0.094	-0.103	-0.103	-0.103	-0.172	-0.172	-0.172
Implied total effect	0.058	0.130	0.163	0.016	0.099	0.027	0.064	0.258	0.061	0.066	0.422	0.063
se	0.047	0.155	0.128	0.023	0.073	0.043	0.056	0.121	0.087	0.065	0.163	0.102
p-value total effet	0.217	0.402	0.204	0.486	0.176	0.529	0.259	0.036	0.480	0.316	0.012	0.542
Implied spillover effect		0.045	0.036		0.039	0.004		0.135	-0.001		0.191	-0.001
se		0.152	0.036		0.078	0.014		0.127	0.022		0.168	0.023
p-value spillover effect		0.766	0.321		0.619	0.780		0.290	0.967		0.260	0.959

Table A18: Spillover analysis among non-network farmers: diffusion channels

	Knows katiti as feed			Has received training on dairy farming			Has been given seeds for free			Has been given seeds by other farmer		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
T	0.015	0.043	0.072**	0.264***	0.208**	0.226***	0.081**	0.178***	0.097***	0.170	-0.013	0.048
	(0.028)	(0.053)	(0.028)	(0.039)	(0.081)	(0.048)	(0.032)	(0.060)	(0.035)	(0.134)	(0.229)	(0.208)
NT_{02}		0.047**			-0.029			-0.011			-0.040	
		(0.023)			(0.023)			(0.017)			(0.114)	
NT_{25}		0.004			0.002			-0.002			0.008	
		(0.006)			(0.007)			(0.006)			(0.017)	
$NT_{02} \times T$		-0.034			0.038			-0.013			0.069	
		(0.024)			(0.024)			(0.018)			(0.115)	
$NT_{25} \times T$		-0.003			-0.002			-0.001			-0.001	
		(0.005)			(0.005)			(0.005)			(0.019)	
$T_{02} \times (1 - T)$			0.109**			-0.073			0.032			-0.203
			(0.048)			(0.057)			(0.042)			(0.231)
Observations	607	607	607	607	607	607	607	607	607	93	93	93
Mean outcome in control	0.111	0.111	0.111	0.236	0.236	0.236	0.111	0.111	0.111	0.696	0.696	0.696
Implied total effect	0.010	-0.003	0.067	0.174	0.228	0.136	0.053	0.043	0.070	0.112	0.225	-0.005
se	0.018	0.058	0.023	0.026	0.074	0.038	0.021	0.056	0.028	0.088	0.238	0.170
p-value total effet	0.587	0.962	0.005	0.000	0.003	0.001	0.012	0.441	0.013	0.210	0.351	0.978
Implied spillover effect		-0.031	0.020		0.091	-0.013		-0.074	0.006		0.233	-0.037
se		0.061	0.009		0.077	0.010		0.059	0.008		0.216	0.041
p-value spillover effect		0.614	0.026		0.237	0.206		0.214	0.449		0.286	0.383

Table A19: Training attended with potential FTs from own and neighboring village

	All farmers		Non network farmers		All farmers	
	Attended training	Number training	Attended training	Number training	Attended training	Number training
Potential FT received training ($Tnei$)	-0.001	-0.002	-0.008**	-0.002	-0.003	-0.003
	(0.003)	(0.005)	(0.003)	(0.006)	(0.004)	(0.006)
Potential FT has signpost ($Spnei$)					0.003	0.003
					(0.005)	(0.007)
Potential FT is in same village ($samevil$)	0.185***	0.491***	0.178***	0.403***	0.185***	0.491***
	(0.017)	(0.065)	(0.036)	(0.104)	(0.017)	(0.065)
Farmer is in treated village (T)	0.002	0.025	-0.003	0.026	0.002	0.025
	(0.004)	(0.018)	(0.006)	(0.021)	(0.004)	(0.018)
$T \times Tnei$	0.023***	0.031	0.033***	0.029	0.020***	0.025
	(0.005)	(0.019)	(0.006)	(0.018)	(0.007)	(0.020)
$T \times Spnei$					0.005	0.012
					(0.007)	(0.013)
$samevil \times T$	0.260***	1.130***	0.207***	0.897***	0.238***	1.172***
	(0.027)	(0.162)	(0.047)	(0.208)	(0.032)	(0.233)
$samevil \times SP$					0.044	-0.084
					(0.027)	(0.223)
Observations (farmer × potential FT pairs)	31,623	31,623	7,673	7,673	31,623	31,623
Mean in control	0.062	0.159	0.051	0.133	0.062	0.159

Notes. Each column displays a separate regression. All regressions include 25 dummies corresponding to the rank of the distance between the FT and the farmer. They also includes lottery strata indicators. Standard errors are clustered at the parish level.