Incentivizing quality in dairy value chains - experimental evidence from Uganda (registered report)

Richard Ariong[*], Jordan Chamberlin[†], Sarah Wairimu Kariuki[†], Bjorn Van Campenhout[‡]

December 4, 2024

Abstract

In value chains where quality of the underlying commodity is hard to observe and track, quality upgrading my be challenging. We test two barriers to the development of a market for quality in Ugandan dairy value chains using a field experiment with treatments at different levels. At the farmer level, we conjecture that farmers are paying attention to the wrong quality attributes and design a video-based information campaign to point out the quality parameters that matter for processors. We also provide them with a small incentive to put what they learned into practice. Midstream, at milk collection centers where milk is bulked and chilled, we provide technology aimed at facilitating quality discovery and tracking. We look at impact of both interventions at both farmer and milk collection center level and consider outcomes such as milk quality, prices received and quantities transacted.

JEL: O13, O17, Q13

Keywords: value chain upgrading, dairy, quality assurance, information

Motivation

Quality of products transacted within value chains, and the preservation of quality throughout the chain, is central to value chain development. Working with quality inputs often reduces production costs further down the value chain. Quality inputs and safeguarding quality while processing, storing, and transporting commodities is also important from a food safety perspective. In

[*]Development Strategy and Governance Division, International Food Policy Research Institute, Kampala, Uganda

[†]CIMMYT, Nairobi, Kenia

[‡]Development Strategy and Governance Division, International Food Policy Research Institute, Leuven, Belgium

general, transformation of value chains often coincide with significant quality upgrading.

Over the past decade, the dairy sub-sector in Uganda has changed dramatically. Particularly in the areas around Mbarara, commonly referred to as the southwestern milk shed, an influx of foreign direct investment has created the preconditions for modern dairy value chains to emerge (Van Campenhout, Minten, and Swinnen, 2021). The area now has an extensive network of milk cooling and collection centers that link smallholder farmers to a cluster of processors. In the dairy value chain, quality is particularly important. Milk quality determines how much end product (eg cheese, caseine, milk powder,..) can be obtained from a given quantity of milk. Furthermore, it goes without saying that the protection of milk from dirt and contamination is important for food safety, as milk is very unstable.

At the same time, it is surprising that there seems to be no market for quality in the Ugandan dairy sub-sector (while this is generally the case in more developed dairy value chain where the price is not a fixed price per liter of milk supplied. Generally, payment differs per dairy farmer, depending on the protein and fat and lactose content of the milk supplied by the farmer[1]). For instance, using recently collected survey data, we find that of a sample of 200 farmers that sold to milk collection centers, only 6 percent indicated that they received a quality premium. From 114 milk collection centers that were included in that survey, we found that only about 18 percent (sometimes) paid a price premium to farmers. At the same time, expert interviews with processors indicate that their main challenge is related to sourcing milk of sufficient quality, pointing out issues related to butter fat content and solid non-fat content of the milk. They also say that the would be willing to pay for it.

When asked about what farmers need to do to increase quality, farmers mainly refer to practices that affect milk sanitation. Most training and extension activities in the area focus on the importance of using proper equipment (stainless steel milk churns as opposed to plastic jerry cans) and simple practices such as washing hands and udders. These technologies and practices do not affect the milk quality attributes that processors seem to care most about. To increase butter fat content and solid non-fat content, it is especially feeding practices that matter.

The above points to at least two problems which constrain the development of a market for quality milk. First, at a technological level, instruments necessary to make the desired quality attributes visible are lacking. Most milk collection centers only engage in rudimentary testing for adulteration (using a gravity based test with a device called a lactometer) and freshness (using the alcohol test). Farmers do not have access to testing equipment. Second, at the knowledge level, farmers do not seem to know what quality parameters are important further downstream the value chain.

In this research, we will test various hypothesis using a randomized control trial with interventions at both the level of the milk collection centers and at

[1] https://www.frieslandcampina.com/owned-by-farmers/milk-price-system/

the farmer level. At the level of the milk collection center, we work with the Uganda Dairy Development Authority (DDA) to scale up their Quality-Based Milk Payment Scheme (QBMPS) that was piloted by last year in Uganda's SW milkshed. It involves installing lactoscans at milk collection centers that allows testing of individual milk deliveries for quality parameters desired by processors. We want to test what the impact of visualizing these quality attributes at this level is on both farmers and milk collection centers. We then use a split plot design to mix in a second intervention at the level of the farmers. Here, we provide a video-based information treatment where farmers are informed about what quality parameters processors deem important and how they can improve on these parameters, and incentivize them by giving a bag of pasture seed.

This document serves as a pre-analysis plan for the study that will be registered in a public repository. It provides background information, outlines hypotheses which will be tested, tools that will be used in the field, power calculations and sample size projections on which sampling is based, outcome variables that will be used to assess impact, and specification that will be estimated. As such, it will provide a useful reference in evaluating the final results of the study (Humphreys, Sanchez de la Sierra, and van der Windt, 2013; Duflo et al., Working Paper).

Related Literature

Our study is related to a large literature. Some of the most recent articles include:

- Rao and Shenoy (2023) explore the effect of collective incentives on group production among rural Indian dairy cooperatives. In a randomized evaluation, they find village-level cooperatives can solve internal collective action problems to improve production quality. However, some village elites decline payments when they cannot control information disclosure. Opting out reflects frictions in allocating surplus within a social network, and suggests some transparency-based efforts to limit elite capture may undermine policy goals.

- Treurniet (2021) uses matching on observable farmer characteristics to study how individual quality incentives provided by private actors can help smallholders to improve milk quality. In the Indonesian dairy value chains they study, individual quality incentives increased the compositional quality of milk quickly after its introduction. Together with physical inputs and training, individual quality incentives also increased the hygienic quality of milk.

- Saenger et al. (2013) use framed field experiment to evaluate the impact of two incentive instruments: a price penalty for low quality and a bonus for consistent high quality milk on farmers' investment in quality-improving inputs among contract farmers in the Vietnamese dairy sector. Statistical analysis suggests that the penalty drives farmers into higher input use, resulting in better output quality. The bonus payment generates even higher quality milk.

Hypotheses and impact pathways

One potential reason why a market for quality does not develop may be related to the fact that milk from individual farmers is poured together, making it hard to track quality. In general, at the start of the cold chain in milk collection centers, only rudimentary testing is done, and equipment to track quality parameters that are most relevant for the development of a market for quality is lacking. Only when milk reaches the processor, these quality parameters are revealed.

In a first hypothesis, we expect that reducing the cost of quality discovery at the level of the milk collection center (such that it is easy to accurately determine the quality of each individual supplier before it is aggregated in milk tanks) will increase outcomes at that level for several reasons. For instance, it will enable collection centers to turn down suppliers with low quality, which should increase the overall quality of milk aggregated. When milk collection centers are able to independently assess the quality of the milk, they may actively search for processors that are prepared to pay a premium for a particular quality parameter.[2] In addition, accurate information about the quality of the milk may also strengthen the bargaining position of the milk collection center vis-a-vis the buyer. The ability to accurately monitor incoming milk may also enable milk collection centers to engage in product differentiation at an early stage, by for instance using one tank to collect high protein milk destined for casseine extraction and using another tank to collect milk that is high in butter fat, to supply to a cheese maker.

In a second hypothesis, we also expect that dairy farmers will benefit from this intervention at the level of the milk collection centers. Making quality visible midstream should enable milk collection centers to reward farmers for supplying superior milk and increase the overall quality of the milk that the collection center aggregates. If dairy farmers know that the milk collection center has the equipment to test milk at a reasonable cost, farmers may also demand milk collection centers to test their milk in case there is discussion related to the quality.

Another potential reason why a market for quality does not develop may be related to the fact that farmers do not have adequate knowledge about what is meant by milk quality. In particular, farmers seem to focus most on food safety related quality aspects of milk, and less on the compositional aspect. As a result, even when the technology to assess quality is available, farmers may not be able

[2]As mentioned earlier, milk quality determines what products can be produced. If the milk collection center discovers their milk has a particularly high butter fat content, it may decide to deliver to a cheese producer who is prepared to pay more for high fat milk than a processor that extracts caseine who is more interested in SNF.

to improve without additional knowledge on what parameters to improve upon. Furthermore, it may be that farmers do not have a good understanding of how these compositional parameters can be affected.[3] A third hypothesis is thus that providing information on what the desired milk quality parameters are, and what affects these parameters, increases outcomes for farmers.

In value chains, it is not always clear whether upgrading is driven by push (eg a productivity increasing technological innovation at the farm level) or pull factors (eg in increase in demand due to opening up of export markets). Often, it is a combination of both, and push and pull factors endogenously reinforce each other in a virtuous cycle (Van Campenhout, Minten, and Swinnen, 2021). In a final hypothesis, we thus also test if making quality visible at the milk collection center level and at the same time providing information on what the desired milk quality parameters are increases outcomes for farmers.

Experimental design

The field experiment consists of two cross-randomized interventions that are implemented at different levels. Outcomes may be measures at different levels. The design is illustrated in Figure 1, which provides a stylized representation of the dairy value chain. We randomly allocate quality testing equipment to a random subset of milk collection centers (MCCs), while another random subset of milk collection centers functions as the control group for this treatment. In the catchment area of each milk collection center, we then take a sample of dairy farmers, stratifying the sample on whether the farmer is an active supplier to the milk collection center or not. In this sample, we then randomly assign half of the farmers to the information treatment (blocking on whether the farmer is an active supplier to the milk collection center or not).

With this design, we can then test if the intervention at the milk collection center improved outcomes for milk collection centers. We can also test if the intervention at the milk collection center affects outcomes at the farmer level by comparing outcomes of the farmers in catchment areas of treated Milk Collection Centers (MCCs) to outcomes of farmers in catchment areas of control MCCs. The intervention at the farmer level can only be evaluated at the farmer level. At the level of the farmers, we can also look at the interaction between the two treatments by looking at outcomes of farmers that received the information treatment in catchment areas of milk collection centers that also received a lactoscan in relation to outcomes of farmers that are differently exposed to the treatments.

In sum, and in reference to the equation we will estimate in the next section, the four main hypotheses that we will test with this design are:

[3]Being a non-rival good, information is generally undersuplied by the private sector. Agricultural extension and advisory services are therefore often organized by governments or non-governmental organizations who tend to prioritize food safety concerns over profitability. As a result, farmers are mostly trained on how to maintain milk sanitary standards and less on ways to improve quality in terms of butter fat and Solid Non-Fat.

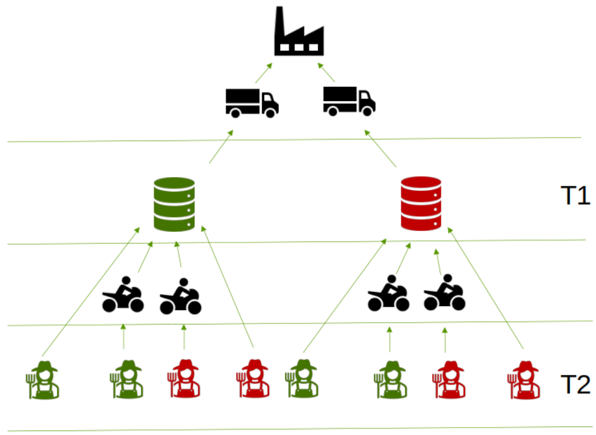

Figure 1: Design

- Hypothesis 1: making quality visible at the MCC level increases outcomes for the milk collection centers ($\beta_{H1} > 0$).
- Hypothesis 2: making quality visible at the MCC level increases outcomes for the farmers in the catchment areas of these MCCs ($\beta_{H2} > 0$).
- Hypothesis 3: providing information on what the desired milk quality parameters are and what affects this parameter increases outcomes for farmers ($\beta_{H3} > 0$).
- Hypothesis 4: making quality visible at the MCC level and providing information on what the desired milk quality parameters are to farmers increases outcomes for farmers ($\beta_{H4} > 0$).

Additional research questions, based on the stratification, tests for differences in average treatment effects between farmers that are connected to milk collection centers versus those that are not. Testing for this treatment heterogeneity allows us to explore if the interventions only strengthen existing value chains or whether they can also draw in actors from informal value chains.

- Does the MCC level intervention affect farmers that are already connected to the milk collection center differently than farmers that are not already connected to an MCC ($\beta_{H2C} = \beta_{H2}$).
- Does the information treatment affect farmers that are connected to an MCC differently than farmers that are not connected to an MCC ($\beta_{H3C} = \beta_{H3}$)
- Does the combined treatment (making quality visible at the MCC level and providing farmers with information on the desired quality dimension)

affect farmers that are connected to an MCC differently than farmers that are not connected to an MCC ($\beta_{H4C} = \beta_{H4}$)

Interventions

To make relevant quality parameters visible at the level of the milk collection centers, we focus on a technology bundle. In close collaboration with DDA, we install milk analyzers at a random sample of milk collection centers. These can be used to test milk samples of individual farmers or traders that supply to the milk collection centers to establish quality of incoming milk, as well as to test samples from the milk tankers when milk is picked up by traders or processors. Milk analyzers show butter fat, solid non-fats, added water, temperature of milk, protein content, and corrected lactometer coefficient. Taking a sample is non-destructive and takes about 30 to 50 seconds depending on the temperature of the milk. Figure 2 shows a milk analyzer during piloting.

The milk analyzers will be delivered with clear Standard Operating Procedure advising MCCs and MCC staff will be trained. We collaborate with the DDA to set up a system to monitor the milk analyzers and its use. In particular, DDA technicians will visit treatment MCCs at set periods. We also set up a hotline that MCCs can contact in case of problems with the milk analyzers. We also make sure that, over the course of the project, equipment is adequately cleaned and calibrated.

In addition to the milk analyzers, the MCC level treatment also consists of a digitized system to keep track of milk quantity and quality delivered to the MCC. To do so, we developed a custom Android application that MCCs can use to register farmers that deliver milk. For these farmers, MCC managers can then record milk deliveries, including quantities delivered and price agreed, as well as a range of quality parameters that can be read from milk analyzer, such as butter fat and protein content. The application can also provide MCC managers with simple reports, such as the average butter fat (weighted by quantities supplied) over a different period (today, yesterday, last week, last two weeks and custom data range). Reports by farmer are also possible, such that MCC managers can determine the total sum to be paid to a farmer for milk delivered in the last 14 days. The application, which is pre-installed on a Samsung galaxy tab A7 with sim-card for mobile internet, backs up data in the cloud, but is designed following an off-line first principle as some MCCs may not have coverage. A screenshot of the application can be found in Figure 3.

Finally, for the MCC intervention, we also developed a poster to be displayed at MCCs informing farmers that the MCC now has a milk analyzer that can determine milk quality for free. The poster was designed by a local artist.

To provide information to dairy farmers on the parameters and characteristics that processors are looking for and how farmers can produce milk that adheres to these standards, we use a short engaging video that demonstrates the inputs and practices that can be used to increase milk quality. The use of video has been found to increase technology adoption in different settings, al-

Figure 2: Milk analyzer

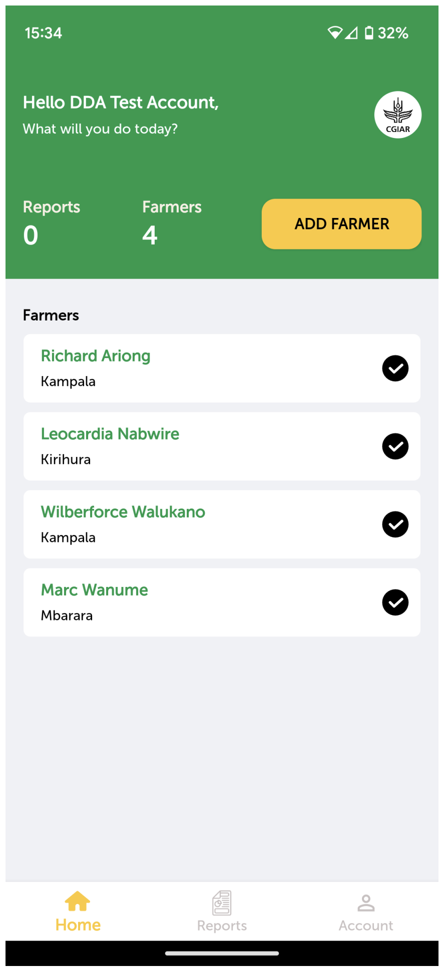

Figure 3: Application

though the effectiveness also depends on a range of design attributes (Spielman et al., 2021). The ability to depict role models in videos seems important to increase both aspirations of the person targeted, as well as creating an enabling environment for adoption in that it may challenge world views and stereotypical thinking (Riley, 2019; Lecoutere, Spielman, and Van Campenhout, 2023).

To design the video based extension intervention, we first identified the top five practices and inputs that are known to raise butter fat and Solid Non Fats in milk. This was done through consultations of experts. We found the top 5 practices and inputs were: selection of breed and genetic potential, selection of grasses for high-quality forage, best practice in silage and hay making, correct mixing and dosage of feed, and feed supplements like Methionine and Lysine. To make the information intervention more actionable, we also provide farmers with some free inputs (1 kg of Chloris Gayana also known as Rhodes grass). The video will be screened a first time during baseline data collection and a second time just before the distribution of the milk analyzers. We also developed an appealing handout that summarizes some of the main points from the video using cartoons drawn by a local artist.

Estimation and inference

We will estimate two equations using Ordinary Least Squares. One equation is at level of the milk collection centers, the second equation is at the level of the dairy farmers.

For the equation that measures impact at the MCC level, denote milk collection centers by m, running from 1 to M. T_m is a treatment indicator at the MCC level that is one if the MCC was allocated to the lactoscan treatment. y_m is the outcome at the level of the milk collection center you want to estimate the treatment effect for and ε_m is an error term. We then estimate the following equation using Ordinary Least Squares. In all equations we also control for baseline outcome if information was collected at that time y_m^b, which will be included in deviations from its mean (Lin, 2013).

$$y_m = \alpha + \beta_{H1}.T_m + \beta_b.y_m^b + \varepsilon_m \tag{1}$$

The parameter of interest in this equation is β_{H1}, which tests Hypothesis 1.

The second equation is at the individual level. Here, T_i is a treatment indicator at the farmer level that is one if the farmer was allocated to the information treatment that informs farmers about what quality parameters are important for processors (with i indicating the farmer running from 1 to I). $C_{i,m}$ is an indicator variable at the farmer level that is one if the farmer i is directly connected to MCC m (and zero if the farmer is connected through a trader). $y_{i,m}$ is the outcome of interest at the level of the individual farmer living in the catchment area of milk collection center m and $\varepsilon_{i,m}$ is an error term (which may be correlated within catchment area). As above, T_m is a treatment indicator at the MCC level that is one if the MCC (in who's catchment area the farmer resides) was allocated to the lactoscan treatment and we also control for the lagged baseline outcome $y^b_{i,m}$,again in deviations from its mean.

$$y_{i,m} = \alpha + \alpha_C C_{i,m} + \beta_{H2}.T_m + \beta_{H3} T_i + \beta_{H4} T_i.T_m$$
$$+ \beta_{H2C}.T_m.C_{i,m} + \beta_{H3C} T_i.C_{i,m} + \beta_{H4C} T_i.T_m.C_{i,m} + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \quad (2)$$

Standard errors in equation 2 are clustered at the milk collection level. The parameter of interest in this equation is β_{H2}, which tests Hypothesis 2, β_{H3}, which tests Hypothesis 3 and β_{H4}, which tests for the interaction effect. We also add a full set of interactions with the connection indicator to look at treatment heterogeneity.

Factorial designs have recently been criticized for the proliferation of under-powered studies and replication failure (Muralidharan, Romero, and Wüthrich, 2023). While in the next section we will run power calculations based on models with a complete set of interactions (equation 2), we may still want to try boosting power by pooling observations across the orthogonal treatment in the event that we find a treatment effect that appears smaller than the minimal detectable effect size that we assumed during power calculations. To do so, we will consider the orthogonal treatment as a co-variate we adjust for, and interact the treatment variable with the demeaned orthogonal treatment. This give a more robust version of the treatment estimate that corresponds to the coefficient estimate of the treatment of interest after dropping the interaction with orthogonal treatment.

To account for multiple comparisons, we will us the two methods illustrated in Anderson (2008). The first consists of computing the Benjamini-Krieger-Yekutieli (2006) sharpened q-values for a vector of p-values We will also combine primary outcomes into an index following Anderson (2008), which also guards against the dangers of multiple comparisons (See Section).

Power calculations

We also use simulation to determine sample size. The primary outcome variable that we use in our statistical power calculations is the price of milk.

We start at the level of the milk collection center and assume that at this level, the price at which milk collection centers sell their aggregated milk is normally distributed with mean 1000 UGX per liter and standard deviation of 50 (which is half of what we will assume at the farmer level). From these N observations (with N denoting the number of milk collection centers recruited for our study and hence the first key variable to be determined by the power calculations) we then generate N times n observations. These are the n dairy farmers that are located in the catchment areas of the N milk collection centers. The outcome variable at this level, prices that farmers obtain from milk collection centers, are generated again as random normal, but with the mean the value that was drawn for the MCC the n farmers are connected to, and with a slightly higher standard deviation (100 — since, as the milk is not aggregated yet, extreme values are not yet averaged out). This procedure gives us a total sample with N prices at the MCC level and $N.n$ prices at the farmer level, the latter being clustered at the MCC catchment area level by design.

We assume that the intervention at the level of the milk collection centers leads to an increase in the price of UGX30 per liter. This seems reasonable in light of the fact that processors told that they either pay a 10 percent premium for quality milk, or UGX100 per liter. However, as we assume a pretty narrow distribution of prices, even though this effect is only a 3 percent increase, this is considered a medium to large effect according to Cohen's D. At the level of the farmers, for the intervention at the MCC level, we expect an effect size of UGX40. While this represents a 4.4 percent increase, the larger variance at this level means that according to Cohen's D, this effect is considered small to medium. Finally, at the level of the farmers, the individual level randomization of the information treatment intervention allows us to estimate small effects. For our power simulation, we assumed and effect size of UGX25, which corresponds to a small effect according to Cohen's D. For the interaction, we assume a large effect (UGX50 per liter).

We calculate power for the joint test that the three hypotheses are true at the 5 percent significance level. To do so, we run the exact two regressions from Section and run 1000 simulations for each $n*N$ combination. For each $n*N$ combination, we calculate the share of simulations at which all coefficients of interest in Equations 1 and 2 (β_{H1} to β_{H4}) are significant at the 5 percent level to determine power.

Results of the simulation are summarized in Figure 4. Instead of the usual power curves that plot power against sample size, we obtain a power plane as we determine both the number of clusters (between 100 and 130 MCCs) and the number of farmers per cluster (between 10 and 40 farmers). Power is measured on the z axis and is the proportion of cases (out of the 1000 simulations) in which all three coefficients were found significant at $p<0.05$.

The figure, which can be found as an interactive figure here, shows the trade-off between more clusters and more individuals per cluster. With about 125 MCCs and 20 farmers per cluster we find power just above .80. This corresponds to a sample of 2500. Note that the requirement to detect minimum effect sizes for all three hypotheses simultaneously is very strict. For instance, if we require only one hypothesis to be significant, we obtain power of .99 for a sample with 125 MCCs and 20 farmers. Similarly, if we consider each hypothesis separately, we get power levels of .87 for the MCC level intervention with outcome at the MCC level, .93 for the MCC level intervention with outcome at the farmer level, and .94 for the farmer level intervention with outcome at the farmer level. For the interaction effect, we obtain power of .99.

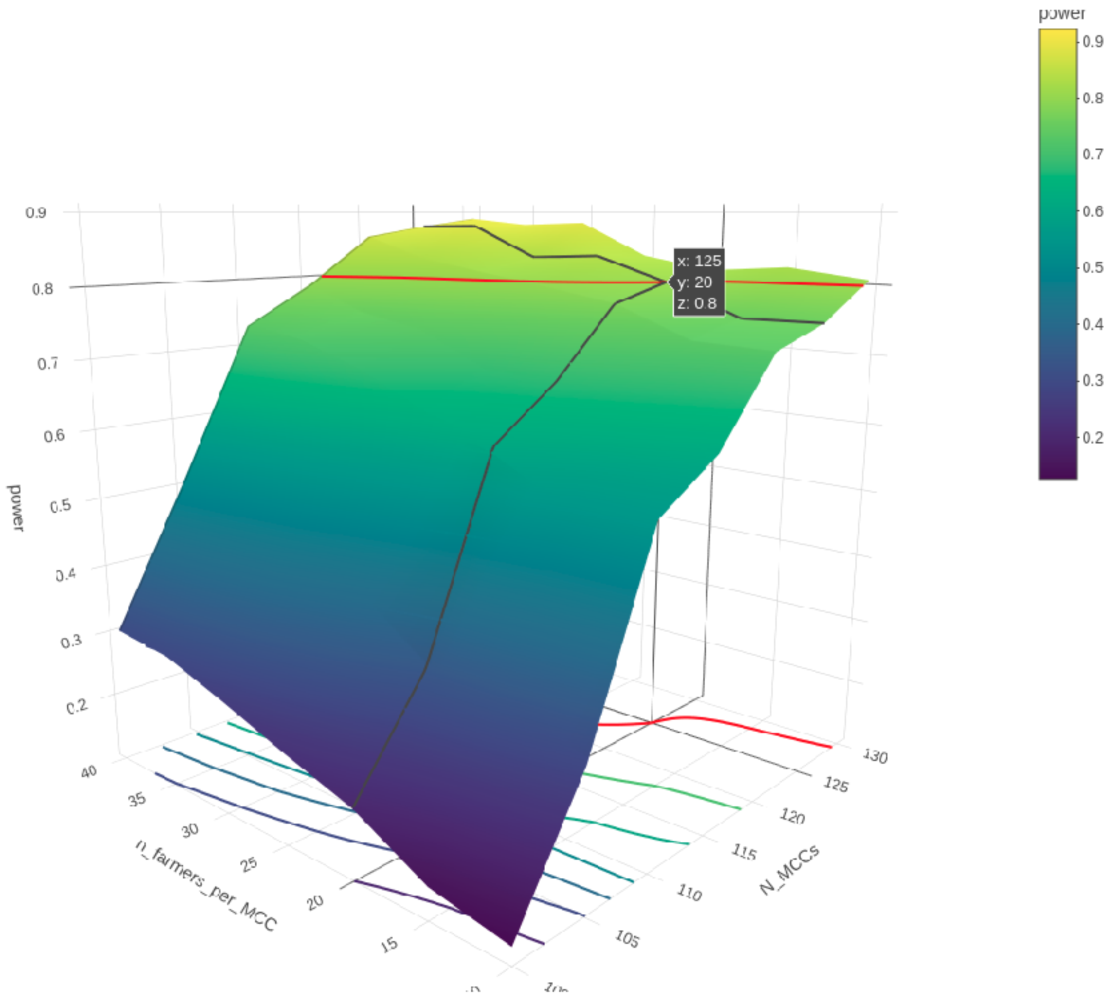

Figure 4: Power plane

Timeline

We plan to collect baseline information in November-December 2022. During that time, we will also implement the intervention at the level of the farmer. Immediately after baseline data collection, we will also start installing lactoscans in the selected milk collection centers. This is expected to take about three months, such that all lactoscans are install towards the end of March 2023. Midline data will be collected about half a year after the last lactoscan was installed, so this will be in September-October 2023. Endline data will be collected one year after the intervention, which is April 2024.

Data collection

Sampling

We start from a list of registered milk collection centers that was obtained from the Dairy Development Authority. From this list, we randomly selected 130 milk collection centers, half of which were assigned to the treatment group using a computer algorithm. We then travel to these 130 milk collection centers and use systematic sampling to get a sample of 10 farmers that are directly delivering to the MCC and 10 farmers that are delivering through a trader to the MCC. In particular, we will visit the MCC early in the morning and get an estimate of the expected number of farmers that will visit during the course of the day. This will be used to determine the interval at which farmers will be picked to participate in the study. These farmers will be interviewed at home the next day.

Data collection protocol

Baseline data collection will be fairly straightforward and consist of surveys at the MCC level and at the farmers level. Endline data collection will proceed in two stages. In a first stage, we will reinterview the farmers and MCCs from baseline. In a second stage, we will also measure quality of milk using milk analyzers in both treatment and control. This will be done by stationing an enumerators in MCCs for one day to test and record all samples that are brought in from the time the MCC opens (usually at 7:00) until it closes for morning receptions (usually around noon). Note that this means that we will also have to install milk analyzers in control MCCs. As effects of installing milk analyzers may be very quick, we only have a limited time window in which differences between treatment and control areas can be measured. Furthermore, we will also need to be mindful of anticipatory effects. That is, we will arrive at MCCs unannounced.

Outcomes of interest

Demonstrating balance

We pre-registered 10 variables at each level to demonstrate balance. Part of these—indicated with a star in the original pre-analysis plan—will be reported in the final paper, the rest will be reported in an online appendix for space considerations. At this stage when we prepare a pre-registered report, we can already estimate balance. Results are in Table 1 for the starred variables; the rest of the results are in Appendix Table 12. We also include F-tests for joint orthogonality.

Primary outcomes

We define six primary outcomes at the MCC level. All six are outcomes are expected to change in a positive direction as a result of our intervention at this level. We thus also construct an index following Anderson (2008) that measures overall effect of making quality visible on the development of a market for quality.

The five primary outcomes at MCC level are:

1. Testing on incoming samples - q25x3
2. Testing of outgoing samples 0-q39a/c, q52a/c, q62a/c, q72a/c, q82a/c
3. Average prices at which milk was bought from farmers (during last 7 days) - q25b
4. Price at which milk was sold (in last 7 days) - q36/q49/q59/q69/q79
5. Does the MCC pay a quality premium to suppliers? - q29
6. Did the buyer pay a quality premium? q44/q54/q64/q74/q84 == 1

Results are in Table 2

We also define five primary outcomes at the farmer level and estimate equation 2. Results are in table 3. In appendix table 13, we estimate the equation where we consider the orthogonal treatment as a co-variate we adjust for, and interact the treatment variable with the demeaned orthogonal treatment. This give a more robust version of the treatment estimate. The primary outcome variables we include are:

1. Production investment and management (based on an anderson index of the following six recommended practices to improve milk quality)
   (a) q39 == "Yes" ##oversowing
   (b) q39c == "Yes" ##legume pastures
   (c) q40 in c(1,3) #controlled/zero grazing in dry seaon
   (d) q41 in c(1,3) #controlled/zero grazing in wet seaon

Table 1: Balance table

	mean ctrl	analyzer	video	bundle	nobs
	milk collection centers				
Is this milk collection center (part of a) cooperative? (yes/no)	0.581	-0.07			124
	(0.495)	(0.089)			
Total Capacity of MCC (in liters)	4592.21	1016.948*			124
	(2471.142)	(427.433)			
Does the MCC pay a premium for quality (yes=1)	0.25	-0.032			124
	(0.435)	(0.08)			
Years Experience in MCC	9.389	0.352			123
	(7.841)	(1.567)			
Facilitates supply of acaracides? (yes=1)	0.548	-0.068			124
	(0.5)	(0.092)			
F-statistic		1.393			
p-value		0.232			
	dairy farmers				
Household Head Age (years)	54.142	0.921	0.097	-2.511	2261
	(13.504)	(2.689)	(2.5)	(3.247)	
Current Total herd size (number)	68.453	12.099	-16.648	18.742	1980
	(80.209)	(16.771)	(13.809)	(21.095)	
Number of improved animals in total herd (share)	76.064	-3.046	-3.408	26.829	2261
	(85.234)	(22.038)	(12.549)	(23.145)	
Liters milk sold per day (on average in the rainy season) (liters)	64.341	8.101	-1.574	11.719	2261
	(66.851)	(16.275)	(11.989)	(17.219)	
Average monthly expense (USD) on chemical purchases	71.87	20.027	-13.912	-24.741	904
	(112.644)	(31.396)	(48.264)	(56.203)	
F-statistic		0.28	0.527	1.323	
p-value		0.924	0.756	0.252	

Note: First column reports control group means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels.

Table 2: Primary outcomes at MCC level

	mean ctrl	analyzer	nobs
Testing on incoming samples	0.276	0.539**	118
	(0.451)	(0.078)	
Testing of outgoing samples	0.276	0.341**	118
	(0.451)	(0.087)	
Average prices at which milk was bought from farmers	1064.655	1.628	118
	(110.191)	(16.398)	
Price at which milk was sold (in last 7 days)	1201.293	-0.149	108
	(108.872)	(20.551)	
Does the MCC pay a quality premium to suppliers?	0.19	-0.018	118
	(0.395)	(0.071)	
Did the buyer pay a quality premium?	0.19	0.031	117
	(0.395)	(0.075)	
Index of primary MCC outcomes	-0.106	0.221*	108
	(0.532)	(0.102)	

Note: First column reports control group means (and standard deviations below); Second column shows average treatment effect of the MCC level intervention (and standard error of coefficient estimate below); **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

(e) q42 == "Yes" #pasture conservation

(f) q43 == "Yes" ##supplements

2. Buyer checks for quality during last transaction (lactoscan, lactometer, alcohol test). q58/qx5/qx17/qx29/qx41/qx53 == 4

3. Price received for milk sold (inclusive of any quality premium that may have been obtained) (price per liter, average of price during last transaction with in last 7 days with buyer) - q55/qx2/qx14qx26/qx38/qx50

4. Does the buyer pay for higher quality milk - q61/qx8/qx20/qx32/qx44/qx56

5. Bargaining power - q65, qx12qx24, qx36, qx48, qx60 ==1 or 3

Finally, we will use data from supervised testing of incoming samples in treatment and control MCCs. In particular, we will test all incoming samples during an entire day. Parameters of interest are the ones that are collected in the app (and can be measured through the milk analyzers): added water, butter fat, SNF, protein, corrected lactometer reading. Results are in Table 4. The data obtained from supervised testing also enables and alternative measure for price effects (as price is also collected during supervised sampling) as well as rejection rates.

Qualitative explorations during the course of the project suggests an interesting side effect of introducing milk analyzers: an increase in freshness. This is because skimming of milk becomes easier to detect and so farmers will abolish this. To fat from milk, milk has to rest a bit after milking as the fat floats on top. If skimming is discouraged, farmers are likely to take milk to MCCs faster, leading to fresher milk.

We test this hypothesis by looking at the distribution of the time at which samples are brought in. We expect that in treatment MCCs, milk is brought in earlier than in control MCCs. In other words, the difference between the time when a sample is brought in and the closing time of the MCC (fixed at 14:00 in our analysis) is likely to increase as a result of the treatment. We investigate this graphically (Figure 5) but also test if distributions are different using a KS-test (test statistic: 0.0018833, p-value: 1) as well as tests for a shift in the distribution using a t-test (test statistic: 2.0370035, p-value: 0.041901) and a Mann-Whitney test (test statistic: 1.56932×10^5, p-value: 0.0189344). Finally, we test for first and second order stochastic dominance. A second way in which we will look at freshness is by the alcohol test.

Secondary outcomes

Secondary outcomes at the milk collection center level include:

1. Enumerator: Do you see the poster advertizing the milk analyzer? - poster

2. Enumerator: Do you see a milk analyzer? - machine

Table 3: Primary outcomes at farmer level

	mean	analyzer	video	bundle	nobs
Production investment and management	0.006	-0.018	0.032	0.018	2128
	(0.564)	(0.046)	(0.037)	(0.053)	
		[0.891]	[0.92]	[0.645]	
Buyer checks for quality	0.183	0.049	-0.003	0.053	1387
	(0.387)	(0.06)	(0.031)	(0.051)	
		[0.59]	[0.853]	[0.341]	
Price received for milk sold	1024.294	13.19	4.98	-3.099	1355
	(117.626)	(12.849)	(11.566)	(14.021)	
		[0.987]	[0.168]	[0.989]	
Get quality premium	0.07	0.006	0.023	-0.03	1387
	(0.256)	(0.027)	(0.023)	(0.029)	
		[0.228]	[0.074]	[0.124]	
Has bargaining power	0.128	0.042	-0.005	-0.03	1387
	(0.334)	(0.04)	(0.021)	(0.037)	
		[0.091]	[0.951]	[0.645]	
Index of primary farmer outcomes	0.007	0.082	0.046	-0.031	1354
	(0.505)	(0.057)	(0.043)	(0.06)	
		[0.099]	[0.106]	[0.488]	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the interaction of the MCC level and farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 4: Milk quality

	mean ctrl	analyzer	nobs
Butter fat	4.029	-0.076	1082
	(0.577)	(0.063)	
SNF	8.873	-0.134	1082
	(0.556)	(0.122)	
Added Water	1.064	-0.146	1082
	(3.715)	(0.346)	
Protein	3.256	-0.035	1082
	(0.197)	(0.044)	
Corrected lactometer reading	29.079	-0.34	1082
	(2.453)	(0.503)	
Index	0.036	-0.07	1082
	(0.795)	(0.096)	

Note: First column reports control group means (and standard deviations below); Second column shows average treatment effect of the MCC level intervention (and standard error of coefficient estimate below); **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values. Added water enters negatively in the Anderson index.

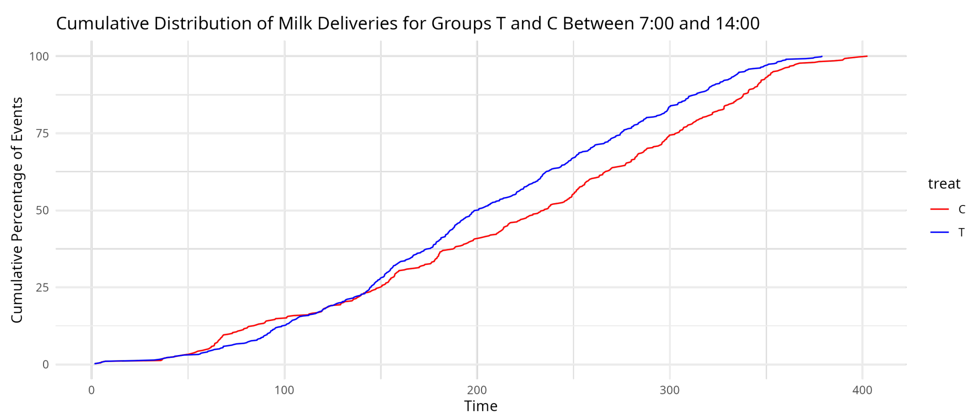

Figure 5: Timing of deliveries

Table 5: Secondary outcomes at MCC level - quantities collected

	mean ctrl	analyzer	nobs
Customers wet season	50.621	6.718	108
	(54.078)	(8.179)	
Customers dry season	47.431	0.651	114
	(56.003)	(9.954)	
Customers last week	49.966	2.011	116
	(55.663)	(6.402)	
Volumes dry season	1384.483	200.569	118
	(927.514)	(182.212)	
Volumes wet season	2521.207	-284.307	118
	(1530.135)	(244.67)	
Volumes last week	2269.086	-256.39	118
	(1536.887)	(233.36)	
Index of secondary MCC outcomes	-0.038	-0.046	106
	(0.783)	(0.107)	

Note: First column reports control group means (and standard deviations below); Second column shows average treatment effect of the MCC level intervention (and standard error of coefficient estimate below); **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

3. Enumerator: Is this the machine that was provided through the project? Make ESSAE - machine_project
4. Enumerator: ask the manager to demonstrate the use of the milk analyzer on the fly and indicate what best maches what transpired machine_in_use==1 or 2
5. How do you keep track of the milk delivered by farmers? record_keeping
6. treatment uptake:
   (a) use of milk analyzer - q16c and q16cc - Information on lactoscan use (for ITT-TOT analysis). - q16

c("tot_sales_q","test_MA","MCC_decides","MCC_got_premium","avg_prem_received")

1. local sales - previous research found that milk collection centers are also important for local milk supply, often doubling as milk shops. Does the intervention crowd out the local market? - q32 == 4X Sold to top 5 processors (Pearl, Amos, Lakeside, GBK, Vital tomosi) (in last 7 days) - q32==2 & q33!=6
2. Xvolumes sold - q35, q48, q58, q68, q78
3. Impact pathway:

Table 6: Secondary outcomes at MCC level - uptake

	mean ctrl	analyzer	nobs
Poster is visible	0.034	0.349**	118
	(0.184)	(0.069)	
Milk Analyzer present	0.224	0.543**	118
	(0.421)	(0.078)	
Project Milk Analyzer is present	0.086	0.63**	118
	(0.283)	(0.07)	
Milk analyzer works	0.224	0.376**	118
	(0.421)	(0.085)	
Milk Analyzer used for almost all incoming samples	0.155	0.395**	118
	(0.365)	(0.081)	
MCC uses App	0.241	0.375**	118
	(0.432)	(0.085)	
Index of MCC uptake	-0.435	0.856**	118
	(0.388)	(0.105)	

Note: First column reports control group means (and standard deviations below); Second column shows average treatment effect of the MCC level intervention (and standard error of coefficient estimate below); **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 7: Secondary outcomes at MCC level - sales

	mean ctrl	analyzer	nobs
Quantity sold	2649.966	-161.447	118
	(2018.357)	(375.337)	
Tested Fat and SNF using MA	0.276	0.333**	118
	(0.451)	(0.087)	
MCC decides	0.052	0.05	118
	(0.223)	(0.049)	
MCC got premium	0.19	0.028	118
	(0.395)	(0.074)	
Average premium received	70.364	-10	7
	(79.915)	(12.472)	
Index of MCC sales	0.391	0.112	7
	(0.545)	(0.172)	

Note: First column reports control group means (and standard deviations below); Second column shows average treatment effect of the MCC level intervention (and standard error of coefficient estimate below); **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

(a) did MCC measure quality of aggregated milk before selling? q37/q50/q60/q70/q80

(b) XIn particular butter fat and SNF using a lactoscan? What equipment was used? q38/q51/q61/q71/q81-q39/q52d/q62d/q72d/q82d

4. Who decided on the price? 1. buyer made offer and MCC accepted, 2. MCC made offer and buyer accepted, 3. negotiation - q40/q53q63/q73/q83 == 2

5. Quality:

   (a) XDid the buyer pay a quality premium? q44/q54/q64/q74/q84 == 1

   (b) How much was the quality premium (UGX per liter)? - q46/q56/q66/q76/q86

   (c) How much passes through to farmers? q31d - share

   (d) How is it distributed to farmers? - distribution

   (e) Does the MCC pay a quality premium to suppliers? - q29

   (f) What was it based on? - q30

6. Does market for quality lead to additional investment in quality preservation - milk cans, etc - q14a-n

7. Does the development of a market for quality lead to more formalization (eg written contracts)

   (a) Between farmer and MCC? (q31)

   (b) Between MCC and buyer? q47/q57/q67/q77/q87

8. Changes in mid-stream service provision: Does the MCC provide services related to

   (a) credit? q17

   (b) Access to acaracides? q18

   (c) Artificial Insemination? q19

   (d) transport? q20

   (e) Training on milk sanitation? q21

   (f) Training on feeding practices? q22

9. product differentiation: do MCCs collect milk in different takers based on quality (eg high protein milk is marketed separately from low protein milk)? - differentiate

10. Buyer changed since start of project? same_buyer

Secondary outcomes at the farmer level:

A first family of secondary outcomes we consider are related to sales by farmers. Results are in Table 8. We asked out average quantities sold on a typical day during the last dry season and during the last rainy season. We also have a binary variable indicating if they sold at any point in the last week. To assess changes at the intensive margin, we also ask about quantities sold during the last sales transactions with each buyer.

Additional family of outcomes looks at choice of buyer:

1. Buyer type sold to on average day in rainy and dry season - q51 and q51x
2. Sold to milk to collection center in the week preceding the survey? (1=yes) - q53==2
3. While we already looked at price received during during transactions is the 7 days prior to endline interview, we also collected data on prices received during the dry and rainy seasons - q51a and q51ax

To assess treatment take-up and compliance, a second family asks weather farmers remember having been shown a video. We also ask if they remember having been given pasture seed and if they do, whether they used it. Finally, we also test of the interventions changed if farmers now start seeing the importance of compositional quality. To do so, we present farmers with three sets of options by asking "If an MCC or processor complains about poor quality milk, which of the 2 options is most important to increase milk quality?" and present two options, one option that focuses on improving milk sanitation (eg washing hands and using milk cans) and one that focuses on compositional quality. (eg using controlled grazing or using feed supplements). The 3 sets of options are then combined and get scored one if the farmer indicates the practice that focuses on compositional quality for all three sets.

A third family of outcomes looks at production, starting with

1. Production (liters) on average day in last wet season - q44 and in the last dry season - q45 and in the last 7 days - q46
2. Home consumption of dairy products (liters (q48), and who consumes diary products - children, calves (q49, q49a)) - test if the development of a market for quality milk crowds out animal sourced food intake within the family or milk as a productive factor
3. Does the intervention affect ghee processing? q66 q68-q69 Does this have gendered effects? q70-q71
4. Main reason for selling to buyer (in particular: because they offer testing, pays premium for quality, payment modalities,...) - q56/qx3/qx15/qx27/qx39/qx51

Table 8: Secondary outcomes at farmer level - Quantities

	mean	analyzer	video	bundle	nobs
Quantity sold in last dry season	58.762	-2.424	2.181	-2.171	2114
	(62.305)	(4.099)	(4.119)	(4.981)	
		[0.641]	[0.9]	[0.856]	
Qantity sold in current season	35.116	-0.833	-0.143	-1.559	2093
	(47.888)	(3.806)	(3.313)	(3.893)	
		[0.279]	[0.475]	[0.599]	
Sold last week?	0.762	-0.022	0.036	0	2217
	(0.426)	(0.039)	(0.026)	(0.04)	
		[0.748]	[0.385]	[0.775]	
Quantity sold in last week	43.798	-4.942	5.97	-3.751	1819
	(74.395)	(5.814)	(6.539)	(7.478)	
		[0.485]	[0.93]	[0.982]	
Index of sales	0.042	-0.093	0.009	0.016	1711
	(0.676)	(0.056)	(0.05)	(0.065)	
		[0.157]	[0.612]	[0.442]	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the interaction of the MCC level and farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 9: Secondary outcomes at farmer level - Sales

	mean	analyzer	video	bundle	nobs
Sold to MCC in wet season	0.706	-0.08	0.03	0.035	2217
	(0.456)	(0.043)	(0.03)	(0.043)	
		[0.935]	[0.104]	[0.89]	
Sold to MCC in dry season	0.699	-0.071	0.04	0.024	2217
	(0.459)	(0.043)	(0.03)	(0.043)	
		[0.923]	[0.053]	[0.961]	
Sold to MCC in last week	0.562	-0.096	0.034	0.044	2217
	(0.496)	(0.051)	(0.034)	(0.049)	
		[0.504]	[0.178]	[0.867]	
Price received in wet season	990.353	9.088	2.22	-18.136	2064
	(153.16)	(14.538)	(9.091)	(13.538)	
		[0.512]	[0.503]	[0.306]	
Price received in dry season	1226.445	8.216	4.029	-6.653	2028
	(173.226)	(15.071)	(13.268)	(16.626)	
		[0.545]	[0.791]	[0.89]	
Index of farmer sales	0.043	-0.056	0.02	0.009	1997
	(0.623)	(0.065)	(0.043)	(0.056)	
		[0.734]	[0.267]	[0.922]	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the interaction of the MCC level and farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 10: Secondary outcomes at farmer level - uptake

	mean	analyzer	video	bundle	nobs
Remembers video	0.356	-0.012	0.322**	0.025	2217
	(0.479)	(0.036)	(0.034)	(0.061)	
		[0.656]	[0]	[0.437]	
Remembers receiving seed	0.424	-0.027	0.347**	0.058	2217
	(0.494)	(0.042)	(0.037)	(0.066)	
		[0.384]	[0.008]	[0.754]	
Used seed	0.284	-0.013	0.231**	0.039	2217
	(0.451)	(0.035)	(0.034)	(0.046)	
		[0.475]	[0.101]	[0.994]	
Knows compositional qualiity matters	0.214	0.049	0.031+	-0.025	2217
	(0.41)	(0.04)	(0.024)	(0.069)	
		[0.408]	[0.305]	[0.969]	
Index of uptake	0	0.032	0.387**	0.018	2217
	(0.638)	(0.049)	(0.047)	(0.084)	
		[0.908]	[0.005]	[0.799]	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the interaction of the MCC level and farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

5. Number of dairy animals (improved/local) - does a market for quality lead to technology adoption for intensification? Is this stronger for the subgroup of farmers that receives the training video, where we explicitly mention that genetics also affect quality parameters? q24-q37

6. Price of dairy animals (improved/local) - test if the development of a market for quality has an impact on the price of animals. q25/q27/q29/q31/q33/q35

7. Gendered decision making outcomes - test if the development of a market for milk impacts who within the household makes the decisions to sell to a particular buyer. q62/qx9/qx21/qx33/qx45/qx57

8. Does the development of a market for quality lead to more formalization and less relational contracting? q63/qx10/qx22/qx34/qx46/qx58

9. Does the intervention also increases milk sanitation (use of milk cans)? q60/qx7/qx19/qx31/qx43/qx55

10. Gendered labour outcomes (milking, marketing, feeding and herding or cleaning)

11. Are farmers aware about the premium offered by buyers? knows_price_downstream/price_downstream

12. Buyer switching behavior. (still_connected==1, q51, q51_prev, q51_name, q51_name_prev) or during dry season (q51x, q51_prevx, q51_namex, q51_name_prevx)?

Ethical clearance

This research received clearance form Makerere's School of Social Sciences Research Ethics Committee (MAKSSREC-10.2022.594/AR) as well as from IFPRI IRB (DSGD-22-1057). The research was also registered at the Ugandan National Commission for Science and Technology (SS1520ES).

Transparency and replicability

To maximize transparency and allow for replicabiliy, we use the following strategies:

- pre-analysis plan: the current document provides an ex-ante step-by-step plan setting out the hypothesis we will test, the intervention we will implement to test these hypotheses, the data that will be collected and specifications we will run to bring the hypotheses to the data. This pre-analysis plan will be pre-registered at the AEA RCT registry.

Table 11: Secondary outcomes at farmer level - switching

	mean	analyzer	video	bundle	nobs
Buyer still connected to baseline MCC?	0.747	-0.094*	0.006	0.03	2126
	(0.435)	(0.044)	(0.023)	(0.035)	
		[0.098]	[0.904]	[0.597]	
Still supplying wet	0.791	-0.108*	-0.01	0.048	1754
	(0.407)	(0.043)	(0.023)	(0.033)	
		[0.06]	[0.426]	[0.924]	
Still supplying dry	0.792	-0.071^{+}	-0.001	0.029	1745
	(0.406)	(0.044)	(0.024)	(0.033)	
		[0.279]	[0.093]	[0.413]	
Index of switching	0.047	-0.23*	-0.017	0.102	1723
	(0.891)	(0.099)	(0.047)	(0.071)	
		[0.088]	[0.306]	[0.969]	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; the third column shows average treatment effect of the interaction of the MCC level and farmer level intervention, (standard error of coefficient estimate below) and [p-value for the hypothesis of no treatment heterogeneity with respect to connection type below]; Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

- revision control: the entire project will be under revision control (that is time stamped track changes) and committed regularly to a public repository (github).

- mock report: After baseline data is collected, a pre-registered report will be produced and added to the AEA RCT registry and GitHub. This report will differ from the pre-analyisis plan in that it already has the tables filled with simulated data (drawn from the baseline). The idea is that after the endline, only minimal changes are necessary (basically connecting a different dataset) to obtain the final result, further reducing the opportunity of specification search.

Funding

This research is part of the OneCG initiative on Rethinking Food Markets which is funded by a consortium of donors. The project builds on previous work on dairy value chains in Uganda that was funded by the Belgian Public Service for Foreign Affairs, Foreign Trade and Development Cooperation, grant number 302015.002.002 (bridge funding). This grant also funded part of the fieldwork for baseline data collection.

Appendix

References

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.

Benjamini, Y., A. M. Krieger, and D. Yekutieli. 2006. "Adaptive linear step-up procedures that control the false discovery rate." *Biometrika* 93 (3): 491–507.

Duflo, E., A. Banerjee, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann. Working Paper. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plans for RCTs in Economics."

Humphreys, M., R. Sanchez de la Sierra, and P. van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21 (1): 1–20.

Lecoutere, E., D. J. Spielman, and B. Van Campenhout. 2023. "Empowering women through targeting information or role models: Evidence from an experiment in agricultural extension in Uganda." *World Development* 167: 106240.

Table 12: Balance table

	mean ctrl	analyzer	video	bundle	nobs
	milk collection centers				
Number of people employed (full-time) at this MCC? (number)	3.145	0.346			124
	(1.987)	(0.333)			
Number of farmers/traders that supply on an average day during the rainy season. (number)	52.339	-6.438			112
	(50.889)	(9.928)			
Capacity use during dry season (share)	35.661	-4.611			119
	(21.66)	(3.669)			
Number of milk cans owned by the MCC	21.452	0.649			124
	(43.62)	(7.762)			
Supplies credit/loans to cooperative members and regularly supplying farmers?(yes=1)	0.863	0.057			124
	(0.345)	(0.069)			
F-statistic		1.791			
p-value		0.121			
	dairy farmers				
Household Members (number)	9.574	-0.064	-0.471	2.096^{+}	2261
	(4.558)	(0.952)	(0.666)	(1.128)	
Liters Produced Total Per Day (average during rainy season) (liters)	76.064	-3.046	-3.408	26.829	2261
	(85.234)	(22.038)	(12.549)	(23.145)	
Normally during the rainy season sells most of its milk to a milk collection center? (yes=1)	0.791	0.027	-0.019	-0.03	2261
	(0.407)	(0.078)	(0.05)	(0.068)	
Uses only steel can/bucket during sales transactions in the last 7 days before survey? (yes=1)	0.778	0.146	0.022	-0.081	2261
	(0.415)	(0.09)	(0.062)	(0.094)	
Member of dairy cooperative? (yes=1)	0.713	-0.191	-0.007	0.083	2261
	(0.452)	(0.119)	(0.058)	(0.09)	
F-statistic		1.353	0.641	1.395	
p-value		0.239	0.668	0.223	

Note: First column reports control group means (and standard deviations below); **, * and + denote significance at the 1, 5 and 10 percent levels.

Table 13: Primary outcomes at farmer level (pooled)

	mean	analyzer	video	nobs
Production investment and management	0.006	-0.044	0.031	2128
	(0.564)	(0.049)	(0.026)	
Buyer checks for quality	0.183	0.058	0.011	1387
	(0.387)	(0.044)	(0.018)	
Price received for milk sold	1024.294	14.897	-10.36	1355
	(117.626)	(9.169)	(5.552)	
Get quality premium	0.07	-0.01	-0.006	1387
	(0.256)	(0.019)	(0.011)	
Has bargaining power	0.128	-0.001	-0.021	1387
	(0.334)	(0.023)	(0.015)	
Index of primary farmer outcomes	0.007	0.046	-0.01	1354
	(0.505)	(0.042)	(0.022)	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below); the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below); Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 14: Secondary outcomes at farmer level (pooled) - Quantities

	mean	analyzer	video	nobs
Quantity sold in last dry season	58.762	-2.317	3.958	2114
	(62.305)	(4.594)	(2.683)	
Quantity sold in current season	35.116	-1.016	1.246	2093
	(47.888)	(3.301)	(2.208)	
Sold last week?	0.762	-0.025	0.026	2217
	(0.426)	(0.029)	(0.016)	
Quantity sold last week	43.798	-2.147	6.878	1819
	(74.395)	(4.473)	(2.83)	
Index of secondary outcomes	0.042	-0.045	0.038	1711
	(0.676)	(0.053)	(0.025)	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below); the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below); Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 15: Secondary outcomes at farmer level (pooled) - Sales

	mean	analyzer	video	nobs
Sold to MCC in wet season	0.706	-0.08	0.02	2217
	(0.456)	(0.038)	(0.019)	
Sold to MCC in dry season	0.699	-0.075	0.021	2217
	(0.459)	(0.037)	(0.018)	
Sold to MCC in last week	0.562	-0.072	0.035	2217
	(0.496)	(0.043)	(0.019)	
Price received in wet season	990.353	-1.657	-8.367	2064
	(153.16)	(12.25)	(5.971)	
Price received in dry season	1226.445	4.888	-2.251	2028
	(173.226)	(13.676)	(7.499)	
Index of sales outcomes	0.043	-0.062	-0.004	1997
	(0.623)	(0.062)	(0.023)	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below); the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below); Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 16: Secondary outcomes at farmer level (pooled) - Uptake

	mean	analyzer	video	nobs
Remembers video	0.356	-0.012	0.321	2217
	(0.479)	(0.03)	(0.02)	
Remembers seed	0.424	-0.003	0.382	2217
	(0.494)	(0.028)	(0.02)	
Used seed	0.284	-0.012	0.244	2217
	(0.451)	(0.027)	(0.018)	
Knows compositional quality matters	0.214	0.036	0.013	2217
	(0.41)	(0.035)	(0.016)	
Index of uptake	0	0.024	0.382**	2217
	(0.638)	(0.04)	(0.026)	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below); the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below); Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Table 17: Secondary outcomes at farmer level (pooled) - Switching

	mean	analyzer	video	nobs
Buyer still connected to baseline MCC	0.747	-0.061	0.017	2126
	(0.435)	(0.036)	(0.016)	
Still supplying wet	0.791	-0.053	0	1754
	(0.407)	(0.037)	(0.018)	
Still supplying dry	0.792	-0.032	-0.006	1745
	(0.406)	(0.037)	(0.018)	
Index of switching	0.047	-0.124	0	1723
	(0.891)	(0.087)	(0.038)	

Note: First column reports control group means (and standard deviations below); the second column shows average treatment effect of MCC level intervention, (standard error of coefficient estimate below); the third column shows average treatment effect of the farmer level intervention, (standard error of coefficient estimate below); Standard errors are clustered at the catchment area level, **, * and + denote significance at the 1, 5 and 10 percent Benjamini-Krieger-Yekutieli sharpened q-values.

Lin, W. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7 (1): 295–318.

Muralidharan, K., M. Romero, and K. Wüthrich. 2023. *Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments*. Tech. rep.

Rao, M. and A. Shenoy. 2023. "Got (clean) milk? Organization, incentives, and management in Indian dairy cooperatives." *Journal of Economic Behavior & Organization* 212: 708–722.

Riley, E. 2019. "Role models in movies: the impact of Queen of Katwe on students' educational attainment." *The Review of Economics and Statistics* 1–48.

Saenger, C., M. Qaim, M. Torero, and A. Viceisza. 2013. "Contract farming and smallholder incentives to produce high quality: experimental evidence from the Vietnamese dairy sector." *Agricultural Economics* 44 (3): 297–308.

Spielman, D., E. Lecoutere, S. Makhija, and B. Van Campenhout. 2021. "Information and communications technology (ICT) and agricultural extension in developing countries." *Annual Review of Resource Economics* 13: 177–201.

Treurniet, M. 2021. "The Potency of Quality Incentives: Evidence from the Indonesian Dairy Value Chain." *American Journal of Agricultural Economics* 103 (5): 1661–1678.

Van Campenhout, B., B. Minten, and J. F. M. Swinnen. 2021. "Leading the way – foreign direct investment and dairy value chain upgrading in Uganda." *Agricultural Economics* 52 (4): 607–631.