Incentivizing quality in dairy value chains - experimental evidence from Uganda

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

September 2, 2025

Abstract

In value chains where quality is hard to observe, upgrading is challenging. We test two barriers in Uganda's dairy sector using a field experiment. At the farmer level, a video campaign addresses misconceptions about quality attributes, paired with small incentives to adopt new practices. At milk collection centers, we introduce technology for testing and tracking. We evaluate both interventions at farmer and center levels, examining impacts on milk quality, prices, and volumes. The results provide evidence on how information, incentives, and technology interact to stimulate the emergence of markets for quality.

JEL: O12, O14, Q13, D82, L15

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

1 Introduction

Quality of products transacted within value chains, and the preservation of that quality throughout the chain, is central to value chain development. High-quality inputs can lower production costs downstream by improving yields, reducing waste, and enhancing processing efficiency. Maintaining quality during processing, storage, and transport is also critical for food safety and consumer health. Beyond these direct effects, consistent quality fosters trust and reputation, supports compliance with domestic and export standards, and facilitates access to higher-value markets. As a result, the transformation of value chains is often accompanied by substantial quality upgrading, reshaping incentives and market structures.

[*]Innovation Policy and Scaling Unit, International Food Policy Research Institute, Kampala, Uganda

[†]CIMMYT, Nairobi, Kenia

[‡]Innovation Policy and Scaling Unit, International Food Policy Research Institute, Leuven, Belgium

For quality to become an important driver of behavior, actors must be properly incentivized to produce and preserve it. In many markets, such incentives take the form of price premiums for higher-quality products, which reward the additional effort or investment required to meet quality standards. When buyers are willing and able to pay more for quality, producers have stronger incentives to adopt better inputs and practices, while intermediaries are more likely to safeguard quality during aggregation, storage, and transport. In the absence of such premiums, quality improvements may not be financially viable, and actors may instead prioritize volume over quality, perpetuating low-quality equilibria.

A central barrier to the emergence of quality premia is the difficulty of observing and verifying quality at the point of transaction. When quality attributes are not directly visible and are costly to test buyers face a classic problem of asymmetric information. Without credible verification mechanisms, high- and low-quality products trade at similar prices, eroding incentives to invest in quality and potentially leading to adverse selection equilibria. This problem is compounded when products from multiple suppliers are pooled before sale, diluting the quality of individual outputs and breaking the link between individual effort and reward. Even when buyers wish to pay for quality, the absence of traceability prevents quality-enhancing investments from translating into higher returns. Further constraints arise from the coordination and investment required to establish quality-based payment systems, including grading, certification, and monitoring infrastructure.

Compounding these challenges is the poor transmission of information about what constitutes quality from the perspective of downstream actors. Absent standards and clear product specifications, producers may hold a different conception of quality than processors, emphasizing attributes visible to them but less relevant to processing outcomes, while overlooking characteristics buyers value most. This misalignment means that effort and resources may be directed toward attributes with little impact on price or marketability, leaving critical quality dimensions unaddressed. Clearly identifying and communicating the specific quality parameters that matter for processors—and providing the means to measure them—can be a high-return investment, enabling price premia to emerge and strengthening incentives for quality upgrading.

The structure of the value chain itself can exacerbate these challenges. When producers and processors are linked only indirectly through layers of traders or informal intermediaries, the likelihood of traceability breaking down increases substantially. At the same time, intermediated chains limit information flows: traders may not transmit, or may even distort, processors' signals about which quality attributes matter most, leaving producers with little guidance for aligning practices with downstream demand (Bray, Serpa, and Colak, 2019). By contrast, in chains where producers are more directly connected to processors, opportunities for credible testing, traceability, and clearer communication of standards are greater, increasing the scope for effective quality-based incentives to emerge.

These challenges are particularly salient in Uganda's emerging dairy subsector. Over the past decade, the industry has undergone rapid transformation, especially in the southwestern milk shed around Mbarara, where foreign direct investment has spurred the development of modern value chains (Van Campenhout et al., 2021). A dense network of milk cooling and collection centers now connects smallholder farmers to a growing cluster of processors. In dairy, quality is central: it determines processing yields for products such as cheese, casein, and milk powder, and maintaining sanitary conditions is essential given the perishability of raw milk. Yet, unlike in more developed dairy markets, there is no functioning market for quality in Uganda. Prices are typically fixed per liter, irrespective of butterfat, solids-not-fat, or protein content. Survey data show that only 6 percent of farmers selling to milk collection centers report receiving a quality premium, and only 18 percent of centers sometimes pay one—even though processors cite insufficient compositional quality as their primary supply constraint and express willingness to pay for improvement.

Part of the problem lies in divergent understandings of quality. Farmers often associate quality with hygiene—clean containers, washed udders—while processors prioritize compositional parameters, which are driven largely by feeding practices. Another part lies in the lack of technological capacity to measure and communicate compositional quality at the point of aggregation: most milk collection centers test only for freshness and adulteration, and farmers rarely have access to test results. Together, these factors weaken the link between effort and reward, suppressing incentives for quality upgrading.

We use this context to test whether the absence of quality premia in Uganda's dairy sector is driven by two constraints: unobservable quality and misaligned perceptions of what constitutes quality. Our experimental design targets these constraints at different points in the value chain. At the midstream level, we work with milk collection centers to make quality observable and recordable in real time through compositional milk testing, digital record-keeping, and farmer-facing information campaigns. At the upstream level, we work with farmers to align their understanding of quality with processor priorities through targeted information, visual reinforcement, and provision of improved pasture seeds to enable quality-enhancing practices.

By implementing these interventions both upstream and midstream, the experiment tests whether reducing information frictions and making quality visible can shift behavior, improve milk quality, and ultimately foster the conditions for a functioning market for quality in the Ugandan dairy value chain. We find that

The paper builds on a large literature that studies how incentives and organizational structures shape quality upgrading in value chains. Rao and Shenoy (2023) demonstrate that group-level incentives in Indian dairy cooperatives can overcome free-rider problems and improve cleanliness, though distributional frictions emerge when elites resist transparency. Treurniet (2021) provides complementary evidence from Indonesia, where individual incentives introduced by processors quickly improved both compositional and hygienic quality of milk, particularly when paired with inputs and training. Saenger et al. (2013) show in the Vietnamese dairy sector that both penalties for low quality and bonuses for high quality increase farmers' investment in quality-enhancing inputs, with bonuses inducing the strongest improvements. Together, these studies highlight the importance of incentive design, the complementarities between price signals and capacity-building, and the risks that organizational frictions pose for sustainability.

This micro-level evidence connects to a broader literature on value chain transformation and quality upgrading in low- and middle-income countries. Bold et al. (2022) show across four field experiments that easing smallholder access to high-quality inputs can promote upgrading, though adoption depends critically on downstream incentives. Fieler, Eslava, and Xu (2018) highlight theoretically and empirically that input-output linkages shape quality upgrading decisions, while Antràs and Chor (2013) emphasizes how the organization of global value chains determines incentives for suppliers to invest in quality. Barrett et al. (2022) situate these findings within the “agrifood value chain revolution,” arguing that the spread of modern procurement systems in developing countries fundamentally reorders incentives, contracting arrangements, and standards.

Our study contributes by examining how technology-enabled monitoring alters the feasibility and credibility of incentive schemes in smallholder-dominated value chains. Whereas Rao and Shenoy focus on the ability of social networks to enforce group incentives, we test whether digital feedback mechanisms can substitute for or complement such informal enforcement. Unlike Treurniet and Saenger et al., who study incentive contracts designed and implemented by processors, we evaluate how introducing quality analyzers at the level of milk collection centers reshapes the space of possible contracts and accountability structures. In doing so, we highlight how technological innovations interact with organizational frictions and market power in ways that may unlock sustained quality upgrading where traditional incentive schemes have struggled.

2 Background: dairy value chains in Uganda

The Ugandan dairy subsector is a dynamic and rapidly growing sector, with its development closely linked to significant policy reforms. Notably, the privatization of the National Dairy Corporation and the enactment of the Dairy Industry Act of 1998 marked critical turning points for the industry. These reforms also led to the establishment of the Dairy Development Authority (DDA), a statutory body under the Ministry of Agriculture, Animal Industry, and Fisheries. The DDA has a dual mandate to regulate and promote the dairy sector, ensuring compliance with quality standards while supporting farmers and stakeholders through training, improved technologies, and market development initiatives.

The policy changes spurred an influx of foreign direct investment, particularly in Mbarara, a key town in southwestern Uganda often referred to as the country’s "milk shed" (Van Campenhout, Minten, and Swinnen, 2021). This investment fostered the emergence of a cluster of milk processors in the region, enhancing value addition and market access. Additionally, productivity in the sector has been bolstered by the widespread adoption of improved dairy breeds, such as Holstein Friesians and Jersey cows, significantly increasing milk yields and overall sector efficiency. Today, the dairy subsector is one of the main export earners of the country.

Dairy value chains in Uganda are diverse and can take various organizational forms, ranging from fully vertically integrated systems involving large processing companies to more fragmented structures dominated by small cooperatives and farmer groups. Despite these variations, a generic value chain consists of five main actors, each playing a distinct role in the production, transportation, and processing of milk.

1. Smallholder Farmers: At the upstream end of the chain are smallholder farmers, who are the primary producers of milk. These farmers deliver milk daily to milk collection centers (MCCs) either personally or by relying on intermediaries such as small traders or transporters.

2. Transporters and Traders: Transporters collect milk directly from farms and deliver it to MCCs for a fee. Unlike transporters, traders purchase milk from farmers with the intent to sell it at MCCs or to other traders for profit, thus functioning as commercial intermediaries within the chain.

3. Milk Collection Centers (MCCs): MCCs are critical nodes in the value chain. Their primary role is to bulk and chill milk, marking the start of the cold chain essential for maintaining milk quality. These centers are strategically distributed across rural areas, facilitating access for smallholder farmers.

4. Large Traders: Once milk is chilled and bulked at MCCs, large traders transport it to processing facilities. This step often involves the use of specialized milk tankers to preserve quality during transit.

5. Processors: Processors, concentrated in key towns such as Mbarara—Uganda's dairy hub—convert raw milk into value-added products with extended shelf lives, such as ultra-heat-treated (UHT) milk, powdered milk, and infant formula. These processors play a pivotal role in integrating Uganda's dairy sector into both domestic and export markets.

This layered and dynamic structure enables the participation of diverse stakeholders while presenting opportunities for efficiency improvements at various points along the chain.

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.[1] 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 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.[2] 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,

[1] 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.

[2] 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.

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.

3 Conceptual framework and simple model

3.1 Environment and timing

There is one processor, many milk collection centers m, and farmers i in each MCC catchment. Quantity per farmer is y_i (treated as given for simplicity), and milk quality is $q_i \in [0, 1]$. Processors pay MCCs a premium for quality when it is verifiable. Two experimental treatments alter information and incentives:

- T1 (MCC level): a measurement-and-transparency bundle at the MCC that makes quality observable at intake and recordable. This raises the premium the processor is willing to pay per unit of quality, and it facilitates pass-through to farmers.
- T2 (farmer level): an information-and-input bundle that improves farmers' knowledge about compositional quality and lowers the cost of taking actions that raise quality.

Within each MCC there are two types of suppliers: those connected directly to the MCC and those who supply via a trader. Traders can capture part of the quality premium.

3.2 Technology and costs at the farmer level

Farmer i chooses effort $e_i \geq 0$ that increases quality:

$$q_i = \theta_i e_i + \varepsilon_i$$

with $\theta_i \geq 0$ the productivity of effort and ε_i a mean-zero shock. Effort cost is convex: $c(e_i) = \frac{k_i}{2} e_i^2$ with $k_i \geq 0$.

Treatment T2 affects the farmer problem in two ways:

it improves knowledge about how actions map into quality, which we capture as an increase in effective θ_i;

it provides inputs that reduce the marginal cost of effort, captured as a decrease in k_i.

3.3 Prices, pass-through, and trader intermediation

Let the processor pay the MCC a unit price

$$P^{\text{proc}} = \bar{p} + \alpha \, \bar{q}_m, \tag{1}$$

where $\bar{q}_m$ is average quality in MCC m and $\alpha \geq 0$ is the per-unit premium slope for verifiable quality. Without verifiable measurement, $\alpha = 0$. Under $T1$, measurement and record-keeping make quality contractible, so $\alpha > 0$.

The MCC offers farmer i a price schedule

$$p_i = p_0 + \lambda_i \alpha q_i, \tag{2}$$

where $0 \leq \lambda_i \leq 1$ is the pass-through rate of the premium to farmer i. Pass-through depends on local contracting frictions and transparency. We assume $\lambda_i = \lambda_C$ if the farmer is directly connected to the MCC and $\lambda_i = \lambda_T$ if the farmer sells via a trader, with $0 \leq \lambda_T \leq \lambda_C \leq 1$. Elements of $T1$ that advertise free testing and create digital records plausibly increase λ_C and may also raise λ_T, but λ_T remains lower because the trader can extract rents.

3.4 Farmer problem and comparative statics

Farmer i chooses effort to maximize

$$\pi_i = \big(p_0 + \lambda_i \alpha q_i\big) y_i - \tfrac{k_i}{2} e_i^2, \tag{3}$$

taking y_i and the price schedule as given. Substituting $q_i = \theta_i e_i$ and ignoring ε_i for the choice, the first-order condition yields

$$e_i^* = \frac{\lambda_i \alpha \theta_i y_i}{k_i}, \tag{4}$$

$$q_i^* = \theta_i e_i^* = \frac{\lambda_i \alpha \theta_i^2 y_i}{k_i}. \tag{5}$$

Comparative statics. For any farmer,

$$\frac{\partial e_i^*}{\partial \alpha} > 0, \qquad \frac{\partial e_i^*}{\partial \lambda_i} > 0, \qquad \frac{\partial e_i^*}{\partial \theta_i} > 0, \qquad \frac{\partial e_i^*}{\partial k_i} < 0. \tag{6}$$

Hence:

- Turning on $T1$ increases α and, via transparency, can increase λ_i. Both raise effort, quality, and farmer outcomes that are increasing in q_i or in the quality-linked component of price.
- Turning on $T2$ increases θ_i and lowers k_i. Both raise effort and quality.
- **Complementarity:**
$$\frac{\partial^2 e_i^*}{\partial \alpha \, \partial \theta_i} = \frac{\lambda_i y_i}{k_i} > 0. \tag{7}$$
The impact of $T2$ is larger when $\alpha > 0$ due to $T1$, and vice versa. The same holds for a decrease in k_i.

Heterogeneity by connection comes from $\lambda_C > \lambda_T$. For given $\alpha, \theta_i, k_i, y_i$,

$$e_{i,C}^* - e_{i,T}^* = \frac{(\lambda_C - \lambda_T)\alpha \theta_i y_i}{k_i} \geq 0, \tag{8}$$

so treatment effects that operate through α and pass-through are larger for directly connected farmers.

3.5 MCC behavior and outcomes

The MCC aggregates milk and sells to the processor. Its per-unit margin is

$$\underbrace{\bar{p} + \alpha \bar{q}_m}_{\text{revenue from processor}} - \underbrace{(p_0 + \lambda_i \alpha q_i)}_{\text{payment to farmer}}. \tag{9}$$

Averaging over suppliers, the expected margin per liter is

$$\text{Margin}_m = \bar{p} - p_0 + \alpha(1 - \bar{\lambda}_m)\,\bar{q}_m, \tag{10}$$

where $\bar{\lambda}_m$ is the average pass-through in the catchment.

Treatment $T1$ raises α and, through farmer responses, raises $\bar{q}_m$. MCC outcomes such as average quality, share of premium contracts, and profits therefore increase under $T1$. If transparency also increases $\bar{\lambda}_m$, the MCC retains a smaller share per unit but benefits from higher volumes of quality and reduced disputes. Net effects on MCC outcomes remain non-negative under mild conditions because both α and $\bar{q}_m$ rise.

3.6 Mapping to hypotheses

Hypothesis 1 (MCC outcomes). Turning on $T1$ increases α and induces higher q_i^*, which raises $\bar{q}_m$ and MCC outcomes that are increasing in realized quality or quality-linked revenue.

$$\text{Prediction: } \beta_{H1} > 0.$$

Hypothesis 2 (farmer outcomes from $T1$). With $\lambda_i > 0$, the increase in α from $T1$ raises e_i^* and q_i^*, and also raises the quality-linked component of the price received. Farmer outcomes that are increasing in quality or in the quality premium rise.

$$\text{Prediction: } \beta_{H2} > 0.$$

Hypothesis 3 (farmer outcomes from $T2$). Treatment $T2$ raises θ_i and lowers k_i, which raises e_i^*, q_i^*, and quality-linked revenues.

$$\text{Prediction: } \beta_{H3} > 0.$$

Hypothesis 4 (combined effect and complementarity). Because e_i^* is multiplicative in α and θ_i/k_i, the cross-partial is positive. The effect of $T2$ is larger when $T1$ is in place and vice versa.

$$\text{Prediction: } \beta_{H4} > 0 \quad \text{for the interaction.}$$

Heterogeneity by connection (C vs. not C). Since $\lambda_C > \lambda_T$, any treatment that works through α has a larger impact for directly connected farmers. Formally,

$$\Delta e^*_{T1|C} > \Delta e^*_{T1|T},$$

and similarly for the combined treatments. This delivers the tests $\beta_{H2C} = \beta_{H2}$, $\beta_{H3C} = \beta_{H3}$, and $\beta_{H4C} = \beta_{H4}$ as equality checks when pass-through is complete within connection type, and suggests treatment effects are larger for connected farmers when $\lambda_C > \lambda_T$.

3.7 Additional testable implications

Beyond the four main hypotheses, the model generates several further predictions:

- Effects scale with baseline volume y_i.
- Effects are stronger for farmers with lower baseline k_i or higher baseline θ_i (or for those who learn the most under $T2$).
- If posters and free testing increase λ_i, treated MCCs should exhibit higher pass-through proxies: more price differentiation by measured quality, fewer disputes, and more frequent test usage.

4 Experimental design

We use a field experiment to test the effectiveness of the innovation bundle. In particular we use a split plot trial design with interventions at two levels of the value chain. The design is illustrated in Figure 1.

The design follows the value chain structure previously outlined, where the first treatment (T1) is implemented at the MCC level. Within each MCC, the second treatment (T2) is randomized at the individual farmer level. Specifically, half of the farmers are assigned to receive the video treatment, while the other half are randomized into the control group. Additionally, within each milk collection center, half of the farmers connected to that MCC receive the video treatment, while the other half are in the control group. To explore heterogeneity in treatment effects, we also stratify by the manner in which farmers are connected to the MCC—either directly or via a trader.

This design allows us to examine outcomes at different levels of the value chain. At the MCC level, we can assess the impact of the T1 intervention on MCC-level outcomes. Furthermore, we can analyze the effect of the second treatment (T2) on farmers connected to the MCC, either directly or through a trader. However, the impact of the farmer-level treatment (T2) can only be evaluated by examining outcomes at the individual farmer level.

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:

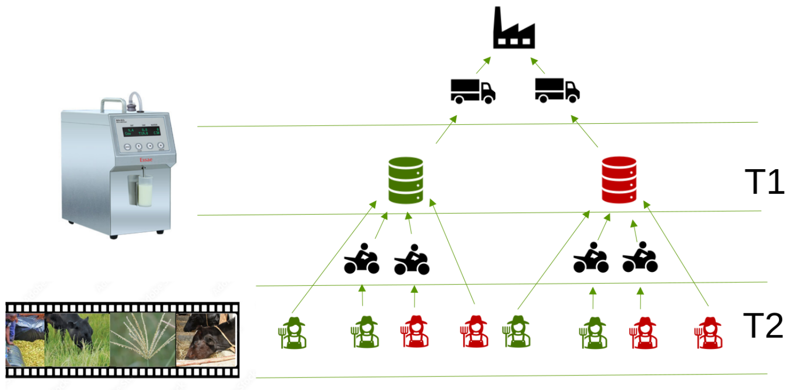

Figure 1: Design

- 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 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 ($\beta_{H2C} \neq \beta_{H2T}$).
- Does the information treatment affect farmers that are connected to an MCC differently than farmers that are not connected to an MCC ($\beta_{H3C} \neq \beta_{H3T}$)
- 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 ($\beta_{H4C} \neq \beta_{H4T}$)

5 Treatments: Innovation Bundles for Addressing Milk Quality Challenges

In response to the challenge of unobservable milk quality in the Ugandan dairy value chain, we designed an innovation bundle targeting Milk Collection Centers (MCCs). This bundle aims to enhance transparency, improve record-keeping, and empower both MCC managers and farmers by making milk quality measurable and visible. The bundle, developed after extensive consultations with stakeholders and implemented together with the DDA, consists of three key components:

1. Milk Analyzer: A central component of the innovation bundle is the installation of a milk analyzer at MCCs. This machine assesses milk quality based on a set of compositional parameters, such as butterfat content and solids-not-fat. The testing process is non-destructive and takes less than one minute per sample, enabling rapid and accurate quality evaluation of all milk collected. By providing immediate feedback, the milk analyzer helps ensure that quality standards are met and maintained. More information on the milk analyzer can be obtained from the manufacturer's website. Figure 2 shows a milk analyzer during piloting.

The milk analyzers were delivered with clear Standard Operating Procedures. Two separate trainings were organized. One training targetted MCC managers, where the focus was mostly on generating buy-in by pointing out the benefits of measuring and tracking milk quality. A second training was geared toward MCC managers and focused more on use and maintanance of the milk analyzers ans

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.

2. IT-Mediated Record-Keeping System: We developed an Android application for MCC managers to facilitate digital record-keeping. The app replaces the traditional paper notebooks used for tracking milk deliveries and payments. In addition to recording quantities and prices, the app allows MCC managers to store and monitor quality parameters obtained from the milk analyzer. It can 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 over a particular time frame, such as in the last 14 days. [3] In addition to the app, we also developed different online portals that can be used to obtain data for different stakeholders. For instance, one portal was geared towards MCC owners, such that they can always monitor key parameters in the MCCs. Another portal aggregated information from all MCCs,

[3] Farmers are typically paid after 14 days of milk delivery.

Figure 2: Milk analyzer

enabling government officials such as the DDA to monitor quality parameters, prices and quantities in real time.[4] This digital system enhances efficiency and transparency, providing both MCC managers and farmers with reliable records that integrate milk quality metrics. The android app can be downloaded from google play store; A screenshot of the application can be found in Figure 3.

Milk collection centers were provided with Samsung Galaxy Tab A7 Android tablet computers on which the application was pre-installed. Each tablet contained a SIM card with a prepaid data bundle to enable cloud-based synchronization of records. The application was developed following offline-first design principles, ensuring that all core functionalities remained accessible without an active internet connection and that data were uploaded automatically once connectivity was restored. Data bundles were topped up by us on a monthly basis throughout the course of the project.

3. Farmer Engagement through Advertisements: To address potential power imbalances emerging from the installation of milk analyzers and ensure farmers are equally informed, we introduced a poster campaign at MCCs. The poster, designed by a local artist, advertises the availability of free milk quality testing for farmers. By making this service widely known, the campaign aims to empower farmers with knowledge about their product quality, fostering trust and collaboration between farmers and MCC managers.

Together, these three components form a comprehensive strategy to make milk quality observable and actionable, creating incentives for improved practices throughout the dairy value chain. This first package will constitute our first treatment in the field experiment below and will be referred to as T1.

To bridge the gap in how milk quality is understood between farmers and processors, we designed an intervention aimed at improving farmers' knowledge of compositional quality and its importance in the dairy value chain. This treatment combines information dissemination, practical advice, and tangible support to drive behavioral changes. It consists of three components:

1. Educational Video: The use of video has been found to increase technology adoption in different settings, although 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). As such, a short, engaging video was developed to explain the concept of compositional milk quality (parameters such as butterfat and protein content) and why it matters for both farmers and processors. The video also highlights practical management practices and inputs that farmers can adopt to improve milk quality, mostly focusing on feeding strategies.[5] The video is designed to be accessible and appealing, ensuring key

[4] The portal can be found here.

[5] To determine the content of the video, 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

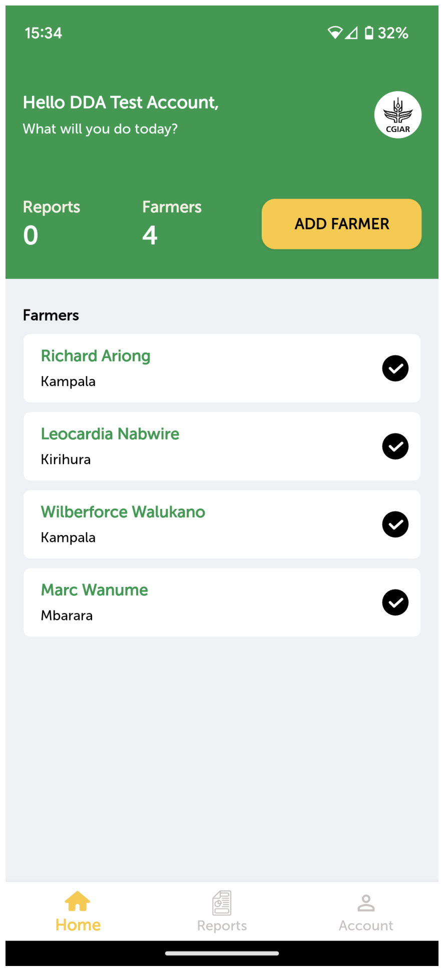

Figure 3: Application

messages resonate with the target audience. The video was shown to treatment farmers twice: once during baseline data collection in December 2022 and once at the time of installation of the milk analyzers in October 2023.

2. Cartoon Handout: To reinforce the information provided in the video, we created a handout in the form of a cartoon summary. The cartoons provide a visual, easy-to-understand recap of the key points, serving as a quick reference for farmers after watching the video. This format ensures the information remains accessible even to farmers with limited literacy.

3. Improved Pasture Seeds: To make the knowledge actionable, each participating farmer received a free bag of improved pasture seeds (1 kg of *Chloris Gayana*, also known as Rhodes grass). By planting these seeds, farmers can enhance the nutritional value of their pastures, a critical factor in increasing the compositional quality of milk. This practical input complements the educational components, enabling farmers to directly apply the recommendations provided.

This treatment, henceforth referred to as T2, aims to align the understanding of milk quality between farmers and processors while equipping farmers with the tools and knowledge to improve milk quality in a sustainable manner.

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^T is a treatment indicator at the MCC level that is one if the MCC was allocated to the test and tracing 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 $\left(y_m^b\right)$.

$$y_m = \alpha + \beta_{H1}.T_m^T + \beta_b.y_m^b + \varepsilon_m \tag{11}$$

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

All other equations use farmer level data. We start by estimating the main treatment effect for T1, the installation of a milk analyzer, on farmers that are associated to the MCC where the milk analyzer was installed. One way to do this is to estimate a fully interacted model such as the one given in Equation 12. Here, T_i^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).

hay making, correct mixing and dosage of feed, and feed supplements like Methionine and Lysine. As selection of breed and genetic potential is unlikely to change sufficiently fast give the length of our research project, we decided to focus on feeding practices.

$$y_{i,m} = \alpha + \beta^*_{H2}.T^T_{i,m} + \beta^*_{H3}.T^I_{i,m} + \beta^*_{H4}.T^T_{i,m}.T^I_{i,m} + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \tag{12}$$

$y_{i,m}$ 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, $T^T_{i,m}$ is a treatment indicator that is one if the MCC (in who's catchment area farmer i resides) was allocated to the milk analyzer treatment and we also control for the lagged baseline outcome $y^b_{i,m}$,again in deviations from its mean. The parameter of interest in this equation would be β^*_{H2}, which tests Hypothesis 2, β^*_{H3}, which tests Hypothesis 3 and β^*_{H4}, which tests for the interaction effect.

While our power calculations relied on models that include the full set of interactions (and we will likewise estimate such a fully interacted specification in Equation 16), an important feature of split-plot (factorial) designs is that power can be increased by pooling observations across the orthogonal treatment. In our case, this means that the effect of the farmer-level intervention can be estimated by pooling across MCC-level treatment arms, and conversely the effect of the MCC-level intervention can be estimated by pooling across farmer-level treatment arms. At the same time, a recent article points out the dangers of pooling treatment cells if there are potential interaction effects between the treatments (Muralidharan, Romero, and Wüthrich, 2023). One way to recover the pooled main treatment effect is to consider the orthogonal treatment as a co-variate and adjust for it, entering it in the regression demeanded and fully intereacted as in equations 13 and 14. 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.

$$y_{i,m} = \alpha + \beta_{H2}T^T_{i,m} + \beta^I_{H3}\left(T^I_{i,m} - \bar{T^I}\right) + \beta^I_{H4}T^T_{i,m}\left(T^I_{i,m} - \bar{T^I}\right) + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \tag{13}$$

$$y_{i,m} = \alpha + \beta_{H3}T^I_{i,m} + \beta^T_{H2}\left(T^T_{i,m} - \bar{T^T}\right) + \beta^T_{H4}T^I_{i,m}\left(T^T_{i,m} - \bar{T^T}\right) + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \tag{14}$$

Next, note that β^*_{H4} in equation 12 is estimated as an incremental effect on top of the main effects. In our analysis, we will test hypothesis 4 by directly comparing farmers that received the training and are connected to an MCC that received a milk analyzer to farmers that did not get the training and are connected to a control MCC. In other words, we will estimate:

$$y_{i,m} = \alpha + \beta_{H4}.T^T_{i,m}.T^I_{i,m} + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \tag{15}$$

after dropping from our data farmers that were exposed to only a single treatment.

Finally, we will also look at treatment heterogeneity, exploiting the fact that we stratified farmers on their link to the MCC. In particular, $C_{i,m}$ 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). We also add a full set of interactions with this connection indicator and run the following model:

$$y_{i,m} = \alpha + \alpha_C C_{i,m} + \beta_{H2T}.T^T_m + \beta_{H3T} T^I_i + \beta_{H4T} T^I_i.T^T_m$$
$$+ \beta_{H2C}.T^T_m.C_{i,m} + \beta_{H3C} T^I_i.C_{i,m} + \beta_{H4C} T^I_i.T^T_m.C_{i,m} + \beta_b.y^b_{i,m} + \varepsilon_{i,m} \quad (16)$$

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. In all regressions, we apply a cluster-robust variance estimator with the bias-reduced linearization (CR2) small-sample correction (Imbens and Kolesár, 2016), with standard errors clustered at the level of randomization (MCC catchment area level).

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 9.1).

6 Sample and Timeline

Sample size was determined using a series of power simulations detailed in the pre-analysis plan. The primary outcome for these calculations was the price of milk, modeled at both the milk collection center (MCC) and farmer level. At the MCC level, prices were assumed normally distributed with mean 1000 UGX/liter and standard deviation 50, while farmer-level prices were drawn with the MCC mean and a higher variance (SD = 100) to capture greater dispersion. We assumed the intervention would increase MCC prices by 30 UGX/liter (medium to large effect size), translate into a 40 UGX/liter increase for farmers (small to medium effect), and that the individual-level information treatment would generate a 25 UGX/liter increase (small effect), with a large interaction effect of 50 UGX/liter. Power was calculated for the joint test of all hypotheses at the 5 percent significance level using 1,000 simulations across varying combinations of clusters (100–130 MCCs) and farmers per cluster (10–40). Results showed that with about 125 MCCs and 20 farmers per MCC (total sample of about 2,500 farmers), power was just above 0.80 for a joint significance test, while power for individual hypotheses was substantially higher, ranging from 0.87 to 0.99. At the design stage, we decided to target 130 MCCs to account for attrition.

The field experiment was conducted in four districts of Southwestern Uganda: Ntungamo, Mbarara, Kazo, and Kirihura. The study began with a comprehensive census of all milk collection centers (MCCs) in the region, from which we sampled 130 MCCs. Half of these were randomly selected to receive

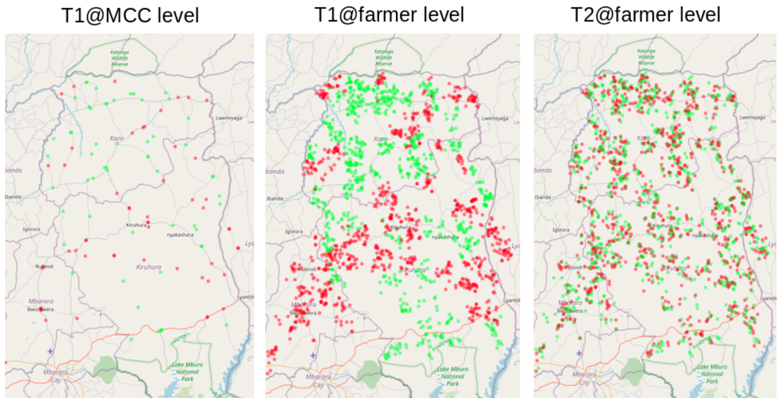

Figure 4: Sampling

T1, while the other half served as the control. In each of the 130 MCCs, we then randomly selected 20 farmers who deliver milk to the center. Of these 20 farmers, 10 were randomly assigned to the video treatment (T2), while the remaining 10 served as the control group.

Figure 4 shows randomization in three of the four districts for T1 and T2 at both the MCC and farmer level. The first map to the left of the figure shows the locations of the MCCs and their treatment status for T1. The second panel shows the treatment status for T1 for the farmers that are connected to the MCCs, which is clustered at the MCC level as T1 is implemented at that level. The third panel shows treatment allocation of T2 among farmers linked to MCCs, which is at the individual level.

Baseline data from both MCCs and associated farmers was collected towards the end of 2022 using in person surveys. At this time, the intervention at the farmer level (T2) was implemented. In the second half of 2023 milk analyzers were installed and MCC managers were trained on the use of the analyzers and the digital record-keeping system. Towards the end of this project, in October 2023, we reinforced T2 by showing the video again to T2 farmers. Next, during the course of one year, together with the DDA we supported T1 MCCs with regular visits to make sure the equipment was operational, including recalibration after 6 months. In december 2024, endline data was collected. This involded in-person surveys at the MCC level and at the farmer level. In addition, during endline we also measured quality of incoming milk. To do so, an enumerator equipped with a milk analyzer paid an unannounced visit to each MCC enrolled in the study and measures all incoming milk samples over the course of one full day.[6] After all data was collected, all equipment was donated

[6]This exercise posed logistical challenges, as MCCs opened early (usually at 7:00 am) and were often in remote areas. Enumerators frequently stayed overnight nearby to set up equipment before deliveries and carried generators for use where power was unavailable. To limit information spillovers, the testing was conducted over a short period with as many

to the MCCs.[7]

7 Descriptive Statistics and Baseline Balance

We pre-registered 10 variables at each level to demonstrate balance. Results for five variables measured at each level (MCC and farmer) are in Table 1; five more variables for each level are in Appendix Table 12 to conserve space (this was also pre-registered). The top panel of Table 1 shows that among MCCs in the control group (i.e., those that did not receive a milk analyzer), 63 percent are organized as cooperatives, reflecting how collective institutions continue to play a central role in the organization of Uganda's dairy value chain. Their average total storage capacity is about 4000 liters, and 27 percent reported paying a premium for higher quality milk. The average MCCs has been in operation for just under 10 years, and 58 percent facilitated the supply of acaracides to their farmer members. The top panel of Appendix Table 12 further shows that among MCCs in the control group, the average number of full-time employees is about three, and a typical MCC receives milk from about 56 farmers or traders on an average day during the rainy season. On average, MCCs use 38 percent of their processing or cooling capacity during the dry season, which is indicative of significant seasonality affecting the sector. MCCs report they own 21 milk cans and the vast majority indicate that they provide credit or loans to cooperative members and regular suppliers.

Both tables further show treatment–control differences for each baseline variable for T1 in the second column. Only one of the individual differences is statistically significant at the 5 percent level (capacity). At the bottom of the top panels of both tables we report results from an omnibus Wald test of joint balance across all covariates (F-statistic and p-values below). These tests do not reject the null of joint equality, indicating that treatment assignment is not systematically predicted by observed baseline characteristics. Overall, the tables suggest that randomization achieved reasonably balanced treatment and control groups along key baseline characteristics of MCCs for the treatment at that level.

The bottom panel of Table 1 reports baseline characteristics of dairy farmers across treatment groups. In the control group (corresponding to farmers that did no see the video and are connected to an MCC that did not receive the milk analyzer), household heads are on average 54 years old, with herds of about 62 animals, of which more than 90 percent are improved breeds.[8] Farmers sell roughly 60 liters of milk per day during the rainy season, and they spend

enumerators deployed simultaneously as possible.

[7]The milk analyzers used by enumerators for supervised testing were donated to control MCCs, which also received tablets and training after the project to ensure fairness.

[8]To calculate herd, we did not simply ask total numbers but asked 6 separate questions: we ask how many local cows; local heifers; and local calves the farmer has, and ask the same 3 categories for improved animals. Farmers were allowed to indicate that they did not know for a particular category, which were treated as missing in our analysis leading to the reduction in the number of observations.

Table 1: Balance table

	mean ctrl	analyzer	video	bundle	nobs
	milk collection centers				
Is this milk collection center (part of a) cooperative? (1=yes)	0.633	-0.086			124
	(0.486)	(0.09)			
Total Capacity of milk tanks (in liters)	4053.167	1031.115*			124
	(1809.592)	(433.343)			
Does the MCC pay a premium for quality (1=yes)	0.267	-0.029			123
	(0.446)	(0.082)			
MCC age in years of operation	9.325	0.235			123
	(7.635)	(1.588)			
Facilitates supply of acaracides? (1=yes)	0.583	-0.068			124
	(0.497)	(0.093)			
F-statistic		1.93			
p-value		0.113			
	dairy farmers				
Household Head Age (years)	54.469	-0.052	0.213	-1.562	2261
	(12.633)	(0.819)	(0.745)	(1.032)	
Current Total herd size (number)	62.297	4.983	5.787	-4.691	1976
	(54.007)	(6.127)	(3.812)	(5.589)	
Number of improved animals in total herd (share)	0.932	-0.007	0.009	-0.014	1976
	(0.171)	(0.012)	(0.011)	(0.014)	
Liters milk sold per day (on average in the rainy season) (liters)	59.697	5.821	7.306*	-7.741	2261
	(56.325)	(6.051)	(3.568)	(4.877)	
Average monthly expense (USD) on chemical purchases	65.551	12.076	15.349+	-29.131*	904
	(91.1)	(11.481)	(7.922)	(13.696)	
F-statistic		0.268	0.497	1.241	
p-value		0.927	0.776	0.314	

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

on average USD 66 per month on chemical purchases—primarily acaracides. Further characteristics in the bottom panel of Appendix Table 12 shows that the average household counts nearly ten members and produces about 70 liters of milk per day during the rainy season. Roughly 79 percent of farmers sell most of their milk to a collection center, three-quarters report using only steel containers when transacting, and a similarly high share are members of a dairy cooperative.

The farmer-level balance tables report three different sets of comparisons. First, we test for pre-treatment differences between farmers connected to MCCs that received a milk analyzer and those connected to control MCCs. With the exception of a modest difference in cooperative membership (significant at the 10 percent level in Appendix Table 12), no systematic imbalance is observed. Second, we examine differences between farmers who received the video-and-seed intervention and those who did not. Here we find some evidence of higher milk production ($p < 0.1$), higher sales volumes ($p < 0.05$), and slightly higher acaracide expenditures among treated farmers. Third, we report the interaction effects at baseline, where two variables are significant at the 5 percent level and one at the 1 percent level. While these individual differences suggest some imbalance across treatment arms, the omnibus F-tests are consistently insignificant.

8 Attrition

Meeting sample size targets within the available budget proved challenging at baseline, largely because dairy farmers are geographically dispersed. In the end, we surveyed 2,261 farmers across 124 MCCs, corresponding to 87 and 95 percent of our planned farmer and MCC samples, and 90 and 99 percent of the required sizes from the power calculations. Of these 124 MCCs, 60 were assigned to control condition and 64 to treatment condition of T1.

By endline, we managed to re-contact 122 MCCs, of which 2 MCCs refused to cooperate, leaving us with data for 120 MCCs. We further managed to contact 2,139, from which 6 refused to be interviewed, leading to an effective sample of 2,133. Further restricting farmers to the subset of farmers connected to an MCC that we were able to interview at endline, we get 2,059.

9 Results

9.1 Primary outcomes

9.1.1 MCC level

We pre-registered six primary outcomes at the level of the MCC. All six are hypothesized to move in a positive direction in response to our intervention, which was designed to make milk quality more transparent and salient in transactions.

To capture the aggregate effect across these outcomes, we also construct a summary index following Anderson (2008). This index provides a single measure of how the intervention influenced the overall development of a market for quality at the MCC level.

Our first two outcomes measure the extent to which quality testing becomes embedded in routine operations at the MCC. Specifically, we record whether MCCs used a milk analyzer to measure butter fat and solid-non-fat content of incoming milk samples in the last 7 days. In addition, we asked whether MCCs test outgoing milk deliveries destined for buyers. This measure is based on more detailed sales transaction data, and is true if for any buyer (eg processor, other trader, etc) butter fat or SNF was tested using a milk analyzer. Together, these indicators capture the degree to which the analyzers are used as intended, and whether testing shapes quality assurance both upstream and downstream.

The next two outcomes reflect how increased transparency affects pricing behavior. We measure the average price at which MCCs purchase milk from farmers during the seven days preceding the survey. We then ask the average price MCCs received for sales to various buyers during the last transaction in the previous week, and take a (weighted) average of these. These outcomes allow us to assess whether improved quality monitoring translates into higher farm-gate prices and whether MCCs are able to cash in on quality improvements through higher sales prices.

Finally, two outcomes directly capture whether price differentiation by quality emerges. We ask whether the MCC pays explicit quality premiums to its suppliers, and whether downstream buyers pay a premium for higher-quality milk. Together, these indicators provide evidence on whether the intervention helped overcome the central coordination problem in quality upgrading: aligning incentives on both sides of the market so that producers and intermediaries are rewarded for investing in quality.

Results are in Table 2. The results show that many MCCs saw the merits of milk analyzers. In particular, MCCs in the treatment group were substantially more likely to test incoming deliveries: the probability of testing rose by 34 percentage points relative to a control group mean of 20 percent, a large and highly significant increase. We also observe a large and significant increase in the probability of testing outgoing deliveries to buyers. These results suggest that the analyzers were effectively used for monitoring both milk supplied by farmers and deliveries to downstream buyers.[9]

Turning to prices, we find little evidence that the intervention affected transaction terms at MCCs. In the control group, the average farm-gate price paid to suppliers was about UGX 1,075 per liter, and we detect no significant increase in the treatment group. Similarly, the average price MCCs received from buyers was about UGX 2,000 per liter in the control group, with no evidence that treatment raised buyer offers. These results suggest that while the intervention influenced testing practices, it did not translate into higher prices on either side

[9]These self-reported outcomes may raise concerns about bias. However, as shown in Table 6, we obtain similar results when using direct enumerator observations of analyzer use.

Table 2: Primary outcomes at MCC level

	mean ctrl	analyzer	nobs
Tested incoming milk using MA (1=yes)	0.203	0.342**	120
	(0.406)	(0.084)	
Testing outgoing milk using MA (1=yes)	0.203	0.285**	120
	(0.406)	(0.084)	
Price at which milk was bought from farmers (UGX)	1075	-15.213	115
	(92.556)	(14.704)	
Price at which milk was sold (UGX)	1199.576	3.439	108
	(106.327)	(19.243)	
Does the MCC pay a quality premium to suppliers?	0.186	-0.032	119
	(0.393)	(0.069)	
Did the buyer pay a quality premium?	0.186	0.031	119
	(0.393)	(0.074)	
Index of primary MCC outcomes	-0.06	0.114	103
	(0.465)	(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.

of the market.

In addition, about 20 percent of MCCs in the control group reported paying a quality premium to suppliers, and the same share (11 out of 59) reported receiving a premium from buyers. Consistent with the price outcomes, we do not find that making quality more visible in the value chain increased the likelihood of such premiums being paid.

Taken together, data recorded at the MCC level suggests that the intervention led to a substantial and consistent increase in testing behavior, but no evidence of quality-based price differentiation. Judged by the composite index of pre-registered MCC-level outcomes, we conclude that T1 had no effect.

9.1.2 Farmer level

We also define four primary outcomes at the farmer level and estimate equation 16, with results reported in Table 3. In Appendix Table 13, we present a robustness specification where the orthogonal treatment is included as a covariate and interacted with the demeaned treatment indicator. This specification provides a more conservative estimate of the treatment effect.

The first outcome measures production investment and management practices that are expected to improve milk quality. We construct an Anderson (2008) index based on six recommended practices: oversowing pastures, planting legume pastures, adopting controlled or zero grazing in both the dry and wet seasons, conserving pastures, and providing feed supplements. This index provides a summary measure of whether farmers adjusted their production strategies in response to the intervention.

The second outcome captures whether buyers actively checked milk quality at the point of transaction. We combine indicators of whether the buyer used a milk analyzer, lactometer, or alcohol test during the most recent purchase. This variable reflects the extent to which farmers were exposed to quality verification in their direct sales.

The third and fourth outcomes focus on prices received. We measure the average price per liter obtained by farmers in their most recent sale within the past seven days, inclusive of any quality premium that may have been awarded. We also directly ask whether the buyer paid more for higher-quality milk. These outcomes allow us to assess whether quality improvements translate into tangible financial benefits for farmers.

Finally, we include an indicator of farmers' bargaining power in transactions with buyers. This measure captures whether the farmer reported being able to negotiate on either price or quality terms, or otherwise exert influence in the transaction. Taken together, these five outcomes allow us to test whether the intervention affected both the upstream adoption of quality-enhancing practices and the downstream rewards farmers received for their efforts.

Table 2 reports treatment effects on our five primary outcomes at the farmer level. Unlike the MCC-level results, where the analyzer intervention led to clear increases in testing (Table 3), the farmer-level results show much weaker impacts. The index of production investments and management practices designed

Table 3: Primary outcomes at farmer level

	mean	analyzer	video	bundle	nobs
Production investment and management (Index)	0	-0.02	0.045	0.044	2054
	(0.56)	(0.039)	(0.024)	(0.052)	
		[0.926]	[0.931]	[0.334]	
Buyer checked for quality (1=yes)	0.186	0.07	0.016	0.108	1337
	(0.39)	(0.047)	(0.021)	(0.058)	
		[0.618]	[0.968]	[0.078]	
Price received for milk sold	1021.887	11.308	1.44	21.011	1254
	(104.146)	(9.852)	(5.104)	(13.083)	
		[0.553]	[0.199]	[0.065]	
Get quality premium	0.07	-0.011	-0.007	-0.002	1337
	(0.256)	(0.022)	(0.013)	(0.027)	
		[0.234]	[0.104]	[0.348]	
Index of primary farmer outcomes	0.002	0.063	0.039	0.159*	1253
	(0.535)	(0.047)	(0.029)	(0.065)	
		[0.278]	[0.209]	[0.016]	

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.

to improve milk quality is slightly positive for both interventions, but neither estimate is statistically significant. This suggests that farmers did not systematically adjust their production strategies in response to either the analyzer at the MCC or the farmer-focused intervention bundle.

Turning to outcomes linked directly to market incentives, we find similarly muted effects. The likelihood that buyers conducted a quality test at the point of sale is unaffected by either intervention. Likewise, farmers in both treatment arms received prices per liter that were statistically indistinguishable from those obtained by control farmers, and they were no more likely to obtain explicit quality premiums. Finally, there is no evidence that either intervention improved farmers' bargaining power in negotiations with buyers. The summary index of all farmer-level outcomes remains close to zero and insignificant across specifications.

Taken together, these findings highlight an important disconnect between the two levels of intervention. At the MCC level, analyzers clearly changed practices by embedding quality testing into transactions. However, these changes did not translate into stronger incentives or rewards for farmers, even when accompanied by a targeted farmer-level intervention. The absence of measurable improvements upstream underscores the difficulty of transmitting quality incentives through the chain: MCCs may increase testing, but without consistent buyer-side demand for quality and corresponding price differentiation, farmers face little reason to adopt quality-enhancing practices. This pattern points to the central coordination challenge in value chains—quality upgrading requires complementary changes across both intermediaries and producers for incentives to align.

9.1.3 Milk samples

To obtain objective measures of compositional quality, enumerators conducted a supervised testing exercise at both treatment and control MCCs. During a full day of collection, all incoming milk was analyzed on-site using milk analyzers. While logistically demanding, this exercise generated high-quality data on several key parameters that are routinely recorded in the app and can be measured with the analyzers: added water, butterfat, solids-not-fat (SNF), protein, and density (corrected lactometer reading). The supervised testing also allows us to construct alternative outcome measures. Since prices were collected alongside the quality tests, we are able to examine the relationship between compositional quality and farm-gate prices. In addition, we can directly measure rejection rates, providing another indicator of whether quality standards were enforced.

Before the supervised data collection began, all analyzers were verified against a reference device calibrated using the ISO 2442 approved method. This verification ensured that the analyzers produced accurate readings across MCCs and that any observed differences in milk quality reflected underlying variation rather than measurement error.

Table 4 presents treatment effects on objective measures of milk composi-

Table 4: Milk quality

	mean ctrl	analyzer	nobs
Butter fat	3.881	0.114*	2518
	(0.539)	(0.036)	
SNF	8.579	0.081+	2518
	(0.49)	(0.05)	
Added Water	1.626	-0.488+	2518
	(3.682)	(0.213)	
Protein	3.163	0.029+	2518
	(0.194)	(0.022)	
Density (CLR)	27.93	0.435+	2518
	(2.31)	(0.197)	
Index	-0.103	0.209**	2518
	(0.817)	(0.054)	

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.

tional quality obtained from the supervised testing exercise. Across several key parameters, we observe consistent improvements in the quality of milk delivered to MCCs in the treatment group.

Milk from treatment MCCs contained significantly higher levels of butterfat and solids-not-fat (SNF), with increases of 0.11 and 0.08 percentage points respectively relative to control group means of 3.88 and 8.58. Protein content also rose by 0.03 percentage points. While these magnitudes are small in absolute terms, they are meaningful given the tight variation in compositional parameters and the importance of butterfat and protein in determining milk value.

At the same time, the prevalence of adulteration declined. Milk in the treatment group contained significantly less added water, with a reduction of nearly half a percentage point relative to the control mean of 1.63. Density (CLR) increased by 0.44, consistent with lower dilution and higher compositional integrity.

The composite index aggregating all five parameters confirms these patterns: overall milk quality improved significantly in the treatment group, with an index effect of 0.21 standard deviations. Taken together, these results suggest that the introduction of milk analyzers not only increased testing behavior at MCCs but also translated into measurable improvements in the quality of milk supplied, reducing adulteration and enhancing key compositional traits.

9.2 Impact on Freshness

Qualitative explorations during the course of the project suggests an interesting side effect of introducing milk analyzers: an increase in freshness. This is

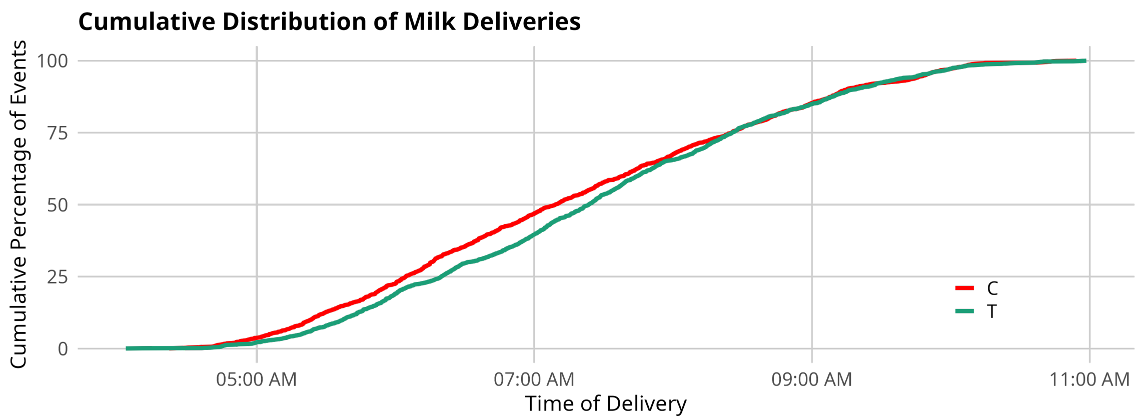

Figure 5: Timing of deliveries

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: 9.6503718×10^{-4}, p-value: 1) as well as tests for a shift in the distribution using a t-test (test statistic: 2.6599864, p-value: 0.0078727) and a Mann-Whitney test (test statistic: 8.08393×10^{5}, p-value: 0.0050529). 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
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

Table 5: Secondary outcomes at MCC level - quantities collected

	mean ctrl	analyzer	nobs
Customers wet season	53.966	4.352	110
	(59.45)	(8.385)	
Customers dry season	48.831	-0.913	116
	(56.549)	(9.881)	
Customers last week	53.102	0.402	118
	(60.21)	(6.472)	
Volumes dry season	1428.814	157.408	120
	(980.508)	(183.545)	
Volumes wet season	2597.119	-347.905	120
	(1625.096)	(244.984)	
Volumes last week	2345.881	-315.412	120
	(1633.783)	(233.945)	
Index of secondary MCC outcomes	-0.018	-0.07	108
	(0.814)	(0.106)	

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.

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 milk analyzer 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:
   (a) did MCC measure quality of aggregated milk before selling? q37/q50/q60/q70/q80
   (b) XIn particular butter fat and SNF using a milk analyzer? What equipment was used? q38/q51/q61/q71/q81-q39/q52d/q62d/q72d/q82d

Table 6: Secondary outcomes at MCC level - uptake

	mean ctrl	analyzer	nobs
Poster is visible	0.034	0.343^{**}	120
	(0.183)	(0.068)	
Milk Analyzer present	0.237	0.533^{**}	120
	(0.429)	(0.078)	
Project Milk Analyzer is present	0.085	0.637^{**}	120
	(0.281)	(0.069)	
Milk analyzer works	0.237	0.369^{**}	120
	(0.429)	(0.084)	
Milk Analyzer used for almost all incoming samples	0.169	0.388^{**}	120
	(0.378)	(0.081)	
MCC uses App	0.237	0.369^{**}	120
	(0.429)	(0.084)	
Index of MCC uptake	-0.432	0.85^{**}	120
	(0.392)	(0.104)	

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	2727.085	-247.552	120
	(2086.725)	(376.422)	
Tested Fat and SNF using MA	0.288	0.328^{**}	120
	(0.457)	(0.086)	
MCC decides	0.068	0.039	120
	(0.254)	(0.051)	
MCC got premium	0.186	0.028	120
	(0.393)	(0.073)	
Average premium received	70.364	-10	7
	(79.915)	(12.472)	
Index of MCC sales	0.403	0.113	7
	(0.532)	(0.171)	

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.

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

Table 8: Secondary outcomes at farmer level - Quantities

	mean	analyzer	video	bundle	nobs
Quantity sold in last dry season	4.281	-0.04	0.05	-0.072	2035
	(1.081)	(0.065)	(0.054)	(0.074)	
		[0.944]	[0.475]	[0.18]	
Qantity sold in current season	3.701	0.018	0.067	-0.121	2019
	(1.111)	(0.08)	(0.057)	(0.081)	
		[0.47]	[0.325]	[0.86]	
Sold last week?	0.797	-0.038	0.013	0.011	2031
	(0.402)	(0.038)	(0.023)	(0.041)	
		[0.552]	[0.79]	[0.553]	
Quantity sold in last week	3.309	-0.274	-0.001	0.128	1601
	(1.938)	(0.188)	(0.145)	(0.197)	
		[0.198]	[0.646]	[0.357]	
Index of sales	0.005	-0.06	0.018	0.004	1579
	(0.657)	(0.053)	(0.04)	(0.068)	
		[0.722]	[0.565]	[0.38]	

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.

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 9: Secondary outcomes at farmer level - Sales

	mean	analyzer	video	bundle	nobs
Sold to MCC in wet season	0.743	-0.07	0.006	0.047	2058
	(0.437)	(0.042)	(0.027)	(0.04)	
		[0.888]	[0.236]	[0.501]	
Sold to MCC in dry season	0.735	-0.061	0.016	0.034	2058
	(0.441)	(0.043)	(0.028)	(0.042)	
		[0.922]	[0.125]	[0.602]	
Sold to MCC in last week	0.588	-0.097	0.014	0.058	2058
	(0.492)	(0.054)	(0.033)	(0.05)	
		[0.491]	[0.379]	[0.61]	
Price received in wet season	990.566	11.592	1.954	-20.465	1994
	(153.302)	(14.35)	(9.105)	(13.507)	
		[0.418]	[0.5]	[0.295]	
Price received in dry season	1225.851	9.895	3.887	-9.879	1958
	(173.874)	(15.184)	(13.236)	(16.621)	
		[0.753]	[0.775]	[0.501]	
Index of farmer sales	0.021	-0.048	0.015	0.013	1929
	(0.625)	(0.063)	(0.041)	(0.056)	
		[0.779]	[0.338]	[0.713]	

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	0.376	0.002	0.326**	0.008	2058
	(0.484)	(0.038)	(0.033)	(0.065)	
		[0.944]	[0.001]	[0.537]	
Remembers receiving seed	0.446	-0.018	0.353**	0.039	2058
	(0.497)	(0.044)	(0.037)	(0.07)	
		[0.576]	[0.012]	[0.682]	
Used seed	0.3	-0.001	0.235**	0.019	2058
	(0.458)	(0.036)	(0.034)	(0.048)	
		[0.787]	[0.102]	[0.903]	
Knows compositional qualiity matters	0.222	0.048	0.024^{+}	-0.029	2058
	(0.416)	(0.042)	(0.024)	(0.073)	
		[0.41]	[0.407]	[0.978]	
Index of uptake	0	0.045	0.382**	-0.008	2058
	(0.63)	(0.05)	(0.045)	(0.084)	
		[0.639]	[0.008]	[0.813]	

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)?

Secondary outcomes at the transaction level:

We should also look at prices here!

10 Heterogeneity with respect to chain length

Our model predicts potential differences in the treatment effect depending on how farmers are connected to MCCs. To test this, we stratified our farmer sample on how they are connected to MCCs: within each MCC, half of the farmers are directly connected, while the other half is connected through an intermediary (see also Figure 1).

To do so, enumerators were required to obtain the list of farmers that actively deliver to the MCC. The list is then split into those that deliver directly, and those that deliver through a trader (either against a fee or a trader working for his/her own account). Systematic sampling was then used to select 10 farmers in each of these two groups: each n-th farmer on the list within each group was interviewed with n being the total number of farmers in the group divided by 10.

Table 11: Secondary outcomes at farmer level - switching

	mean	analyzer	video	bundle	nobs
Buyer still connected to baseline MCC?	0.765	-0.065	0.007	0.033	2052
	(0.424)	(0.041)	(0.023)	(0.036)	
		[0.204]	[0.947]	[0.453]	
Still supplying wet	0.807	-0.081	-0.008	0.054	1692
	(0.394)	(0.04)	(0.023)	(0.034)	
		[0.07]	[0.449]	[0.566]	
Still supplying dry	0.807	-0.044	0.002	0.031	1685
	(0.395)	(0.041)	(0.024)	(0.034)	
		[0.359]	[0.082]	[0.619]	
Index of switching	0.049	-0.163^{+}	-0.011	0.115	1663
	(0.882)	(0.092)	(0.049)	(0.075)	
		[0.14]	[0.283]	[0.675]	

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.

Unfortunately, this procedure did not always work in practice. Particularly in MCCs that take a coopertive form, it was often difficult to find sufficient traders, and so proportionally more directly collected farmers were selected. In particular,

11 Conclusion

This study explored the potential of targeted interventions to accelerate quality upgrading in Uganda's dairy value chain. Using a field experiment, we introduced innovations at the MCC and farmer levels, including milk analyzers, digital record-keeping tools, and farmer training. The interventions aimed to address persistent challenges related to milk quality by making it observable, actionable, and incentivized. This comprehensive design allowed us to assess impacts at multiple levels of the value chain, providing robust insights into what drives—or limits—quality improvements.

Findings:

- analyzers were effectively used for monitoring both milk supplied by farmers and deliveries to downstream buyers - at least that is what MCCs told us - secondary outcomes also suggest MAs are used, a good share of MAs also put up the posters and used the app.
- No clear price effects - which of the following explanations are most plausible
  - Thin margins at MCCs: MCCs typically operate on tight margins, with their sales price closely tied to what processors or other downstream buyers pay them. If buyers did not pay more for higher quality milk, MCCs had little scope to pass quality information through prices. - this seems consistent with the
  - Contracts and price stickiness: Prices may be set in advance or negotiated infrequently (daily/weekly/monthly), making them less responsive to changes in quality testing over the study period. - this is also plausible in the context, where prices are typically fixed for 14 days
  - Market structure: If competition among MCCs is limited, or if buyers wield monopsony power, MCCs may be unable to raise prices for farmers even if quality improves.
  - Heterogeneous effects masked in averages: Rather than raising the average price, quality testing may have led to more price differentiation: better-quality suppliers receiving a premium, but lower-quality suppliers being rejected or paid less. This could leave the mean unchanged even if underlying incentives shifted. - this does not seem to be the reason but we can investigate this using the testing data.

– Adjustment margins other than price: MCCs may have used testing results to enforce non-price mechanisms such as rejecting adulterated milk, reducing disputes, or strengthening supplier relationships.

– Short time horizon: It may take time for buyers to recognize and reward systematic improvements in quality. The intervention might influence quality assurance practices first, with price effects emerging later.

The current price-setting mechanism is a critical bottleneck. Processors continue to publish uniform prices for milk, undermining the alignment of quality improvements with financial rewards. However, recent developments indicate potential shifts in the market. Two of Uganda's largest processors are piloting quality-based payment systems, directly rewarding farmers for higher-quality milk. One MCC SUMPCArecently started selling milk to one of the large processors through their quality based milk payment system. Additionally, a large new processor specializing in infant formula production that is entering the market has already started contacting MCCs to gauge their interest in supplying milk under a quality based payment system.

These findings highlight both the progress made and the challenges that remain. While the interventions successfully improved milk quality and increased transparency, systemic changes—particularly in pricing mechanisms—are essential to sustain and scale these improvements. Future efforts should focus on supporting emerging market trends, addressing adoption barriers, and ensuring that economic incentives align with quality-based practices, ultimately benefiting all stakeholders across the dairy value chain.

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.

- 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.

12 Acknowledgments

During the preparation of this work the author(s) used OpenAI's chatGPT in order to obtain editorial suggestions to improve clarity and readability and explore alternative phrasing for technical terms. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund: https://www.cgiar.org/funders/. In particular, funding from the CGIAR Rethinking Food Markets Initiative and the CGIAR Science program on Better Diets and Nutrition is greatly appreciated. We want to thank Leocardia Nabwire, Wilberfoce Walukano and Marc Charles Wanume for field support.

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.

Antràs, P. and D. Chor. 2013. "Organizing the global value chain." *Econometrica* 81 (6): 2127–2204.

Barrett, C. B., T. Reardon, J. Swinnen, and D. Zilberman. 2022. "Agri-food value chain revolutions in low-and middle-income countries." *Journal of Economic Literature* 60 (4): 1316–1377.

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.

Bold, T., S. Ghisolfi, F. Nsonzi, and J. Svensson. 2022. "Market access and quality upgrading: Evidence from four field experiments." *American Economic Review* 112 (8): 2518–2552.

Bray, R. L., J. C. Serpa, and A. Colak. 2019. "Supply chain proximity and product quality." *Management science* 65 (9): 4079–4099.

Fieler, A. C., M. Eslava, and D. Y. Xu. 2018. "Trade, quality upgrading, and input linkages: Theory and evidence from Colombia." *American economic review* 108 (1): 109–146.

Imbens, G. W. and M. Kolesár. 2016. "Robust Standard Errors in Small Samples: Some Practical Advice." *The Review of Economics and Statistics* 98 (4): 701–712.

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.

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.

Online Appendix

Table 12: Balance table

	mean ctrl	analyzer	video	bundle	nobs
	milk collection centers				
Number of people employed (full-time) at this MCC? (number)	2.967	0.346			124
	(1.886)	(0.337)			
Number of farmers/traders that supply on an average day during the rainy season (number)	55.923	-6.573			112
	(64.037)	(10.032)			
Capacity use during dry season (share)	37.978	-4.564			119
	(20.991)	(3.718)			
Number of milk cans owned by the MCC	21.05	0.716			124
	(52.963)	(7.877)			
Supplies credit/loans to cooperative members and regularly supplying farmers? (1=yes)	0.833	0.057			124
	(0.376)	(0.07)			
F-statistic		1.604			
p-value		0.269			
	dairy farmers				
Household Members (number)	9.754	-0.512	-0.314	0.925^{*}	2261
	(4.687)	(0.342)	(0.252)	(0.357)	
Liters Produced Total Per Day (average during rainy season) (liters)	70.053	9.438	7.726^{+}	-10.447^{+}	2261
	(61.477)	(7.614)	(4.051)	(6.076)	
Normally during the rainy season sells most of its milk to a milk collection center? (1=yes)	0.793	-0.023	0.006	0.025	2261
	(0.405)	(0.049)	(0.015)	(0.028)	
Uses only steel can/bucket during sales transactions in the last 7 days before survey? (1=yes)	0.752	0.046	0.029	-0.046	2261
	(0.432)	(0.039)	(0.019)	(0.03)	
Member of dairy cooperative? (1=yes)	0.757	-0.084^{+}	0.004	-0.009	2261
	(0.429)	(0.048)	(0.017)	(0.028)	
F-statistic		1.283	0.605	1.322	
p-value		0.288	0.696	0.273	

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	-0.02	0.045	2054
	(0.56)	(0.039)	(0.024)	
Buyer checks for quality	0.186	0.07	0.016	1337
	(0.39)	(0.047)	(0.021)	
Price received for milk sold	1021.887	11.308	1.44	1254
	(104.146)	(9.852)	(5.104)	
Get quality premium	0.07	-0.011	-0.007	1337
	(0.256)	(0.022)	(0.013)	
Index of primary farmer outcomes	0.002	0.063^{NA}	0.039^{NA}	1253
	(0.535)	(0.047)	(0.029)	

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	4.281	-0.072	0.069	2035
	(1.081)	(0.104)	(0.045)	
Quantity sold in current season	3.701	-0.059	0.037	2019
	(1.111)	(0.096)	(0.047)	
Sold last week?	0.797	-0.035	0.01	2031
	(0.402)	(0.028)	(0.016)	
Quantity sold last week	3.309	-0.139	0.114	1601
	(1.938)	(0.169)	(0.073)	
Index of secondary outcomes	0.005	-0.079	0.026	1579
	(0.657)	(0.05)	(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 15: Secondary outcomes at farmer level (pooled) - Sales

	mean	analyzer	video	nobs
Sold to MCC in wet season	0.743	-0.074	0.007	2058
	(0.437)	(0.039)	(0.019)	
Sold to MCC in dry season	0.735	-0.07	0.007	2058
	(0.441)	(0.038)	(0.019)	
Sold to MCC in last week	0.588	-0.07	0.026	2058
	(0.492)	(0.046)	(0.021)	
Price received in wet season	990.566	-1.213	-10.037	1994
	(153.302)	(12.508)	(6.003)	
Price received in dry season	1225.851	3.156	-3.15	1958
	(173.874)	(13.979)	(7.626)	
Index of sales outcomes	0.021	-0.059	-0.006	1929
	(0.625)	(0.064)	(0.024)	

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	0.376	-0.005	0.325	2058
	(0.484)	(0.031)	(0.021)	
Remembers seed	0.446	0.003	0.391	2058
	(0.497)	(0.029)	(0.02)	
Used seed	0.3	-0.005	0.25	2058
	(0.458)	(0.028)	(0.019)	
Knows compositional quality matters	0.222	0.036	0.01	2058
	(0.416)	(0.037)	(0.017)	
Index of uptake	0	0.031	0.381^{**}	2058
	(0.63)	(0.04)	(0.027)	

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	0.765	-0.037	0.02	2052
	(0.424)	(0.033)	(0.017)	
Still supplying wet	0.807	-0.029	-0.001	1692
	(0.394)	(0.035)	(0.018)	
Still supplying dry	0.807	-0.007	-0.008	1685
	(0.395)	(0.035)	(0.018)	
Index of switching	0.049	-0.062	-0.003	1663
	(0.882)	(0.082)	(0.04)	

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.