The Impact of Public-Private Partnerships in Food Value Chains in Lower Middle-Income Countries

Mark Treurniet

The Impact of Public-Private Partnerships in Food Value Chains in Lower Middle-Income Countries

Mark Treurniet

Thesis Committee

Promotor

Prof. Dr E.H. Bulte

Professor of Development Economics

Wageningen University & Research

Co-promotor

Dr W.J.J. Bijman

Associate Professor, Business Management & Organization Group

Wageningen University & Research

Other members

Prof. Dr F. Alpízar, Wageningen University & Research

Prof. Dr T. Bernard, International Food Policy Research Institute, Washington DC, USA

Dr B. Dillon, Cornell University, Ithaca NY, USA

Dr M. Torero, Food and Agriculture Organization, Rome, Italy

This research was conducted under the auspices of the Wageningen School of Social Sciences (WASS)

The Impact of Public-Private Partnerships in Food Value Chains in Lower Middle-Income Countries

Mark Treurniet

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 3 June 2020
at 1:30 p.m. in the Aula

Mark Treurniet
The Impact of Public-Private Partnerships in Food Value Chains in Lower Middle-Income Countries, 166 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2020)
With references, with summaries in English and Dutch

ISBN 978-94-6395-331-3

DOI https://doi.org/10.18174/516597

To those that are poorly treated by economic systems

Contents

1. General Introduction

Mark Treurniet

1

1.1. Background

1.1.1. Sustainable Development Goals

Since 1990, the number of people living in extreme poverty has decreased by more than 60%. However, 731 million people are still living on less than $1.90 a day.[1] Most of these people live in lower middle-income countries (World Bank 2019a).[2] The vast majority live in rural areas, and their incomes greatly depend on agricultural activities (De La O Campos et al. 2018).

Meanwhile, ensuring universal access to safe food is a big challenge, particularly in poor areas. While global foodborne hazards cause 600 million illnesses and 420 thousand people deaths, the highest food borne disease burden per capita is observed in Africa, and the second-highest in South-East Asia (WHO 2015).

Both ending poverty and ensuring universal access to safe food are important objectives of the United Nations' Sustainable Development Goals (SDGs). Ending poverty is the core of SDG 1, while ensuring universal access to safe food is an important part of SDG 2.

1.1.2. The rise of Aid for Trade

For a long time, it was believed that tariff reductions are sufficient to increase exports and stimulate development. After decades of multilateral negotiations, the Uruguay Round of 1986-1993 led developing countries to accept the most ambitious trade liberalization in history, which was promised to deliver large welfare gains, from which a large part would accrue to developing countries (Stiglitz and Charlton 2013).

It soon became clear, however, that the agreements were unfair towards developing countries, since (i) market access gains were relatively small for sectors that are important for developing countries, (ii) the ability of poor countries to export was limited by non-tariff barriers, weak infrastructure and supply constraints, and (iii) costs of establishing institutions and regulations were huge (Stiglitz and Charlton 2013).

---

[1] 2011 Purchasing Power Parity.

[2] Lower middle-income countries are a World Bank classification for countries with a GNI per capita, calculated using the World Bank Atlas method, between $996 and $3,895 in 2017.

Moreover, the assumed positive relationship between trade and development has been reappraised over time. First, in a closed economy, domestic production and domestic prices are negatively related. Trade liberalization weakens this negative relation, and thus causes farmer incomes to vary more with production. This induces people to shift away from risky, but more productive technologies (Newbery and Stiglitz 1984), particularly in developing countries with imperfect risk markets. Second, labor does not easily move from inefficient to efficient sectors, due to imperfections in financial markets and lack of entrepreneurship. Third, trade liberalization takes away tariffs as an important source of government revenue, which forces governments to lower public investments in health and education. Especially poor people might be excluded from benefits of trade liberalization (Stiglitz and Charlton 2013).

In 2005, the unfairness of historical trade agreements and the disappointing results from trade liberalization for developing countries, ultimately led to the launch of the Aid for Trade initiative, whose objectives are reflected in the WTO Hong Kong Ministerial Declaration (2005): "Aid for Trade should aim to help developing countries, particularly LDCs,[3] to build the supply-side capacity and trade-related infrastructure that they need to assist them to implement and benefit from WTO Agreements and more broadly to expand their trade."

The Aid for Trade initiative was embraced by the development community, which during the same times struggled to scale up its disbursements and demonstrate its effectiveness (Easterly 2009). Although Aid for Trade was meant to be additional to existing aid, it has basically become a grouping within Official Development Assistance (ODA), and may have crowded out non-trade-related ODA (Razzaque and Velde 2013, Stiglitz and Charlton 2013). The share of Aid for Trade in total ODA has increased over time, as Aid for Trade commitments have grown from an average of $22 billion in 2002-2005 to more than $50 billion in 2015 and 2016, and now amount to approximately 30% of total ODA (OECD/WTO 2017, OECD 2018).

Table 1.1 shows a simple breakdown of current Aid for Trade over its four main categories. Most of Aid for Trade goes to economic infrastructure (59%) and building productive capacity (39%). Within the latter category, most aid flows to food value chains. Since consumers may

[3] Least Developed Countries are a UN classification for countries with low levels of per-capita income and human assets, and high vulnerability to economic and environmental shocks.

want to pay higher prices for safe food, interventions in the food value chains could potentially increase both food safety (SDG 2) and farmer incomes (SDG 1).

Table 1.1. Aid for Trade Commitments by Category for 2015

	USD million	Percent
Economic infrastructure	31,784.3	59.0%
Building productive capacity	21,030.5	39.0%
- Of which in agricultural sector	9,632.7	17.9%
Trade policy and regulations	1,061.8	2.0%
Trade-related adjustment	2.2	0.0%
Total	53,878.8	100.0%

Source: OECD/WTO (2017)

1.1.3. Public-private partnerships in food value chains

Aid for Trade in food value chains often involves the private sector through public-private partnerships (PPPs). A PPP is "a long-term contract between a private party and a government entity, for providing a public asset or service, in which the private party bears significant risk and management responsibility, and remuneration is linked to performance" (World Bank 2019b). While this definition remains vague about "performance" and "risk", remuneration is generally linked to the *outputs* delivered, like for example a road being built or farmers being trained, and the private party bears the risk of the delivery of these outputs. PPPs are therefore conceptually different from Development Impact Bonds (DIBs), in which remuneration is contingent on development *impacts* realized, like for example increased food safety or increased incomes. DIBs are further discussed in the General Discussion of this thesis.

Donors may have various reasons to involve the private sector through PPPs. First, PPPs may help to use the expertise and efficiency of the private sector to realize development goals. Second, private companies may want to leverage public investments by co-funding interventions. Third, as a result of PPP interventions, the donor countries' private sector may benefit from increased access to high-quality inputs and local consumer markets, creating the promise of a win-win situation (Poulton and Macartney 2012, Bitzer and Glasbergen 2015).

Poulton and Macartney (2012) discuss the various market failures that PPPs in food value chains might help to solve. First, public good aspects of information provision might lead to under-investment of outreach to producers and consumers. Second, smallholders may have limited access to technical information, capital, input, risk and output markets, and therefore face high barriers to entry. Third, asymmetric information and lack of commitment devices may cause coordination failure.[4]

1.2. Impact of Aid for Trade

1.2.1. Over-all impact

Despite the huge investments in Aid for Trade, rigorous evidence on the impact of Aid for Trade on exports and development has remained scant. The thin evidence suggests that over-all Aid for Trade increases exports, but that this effect is entirely driven by aid to economic infrastructure. Aid to productive capacity was not found to affect over-all exports (Calì and te Velde 2011, Vijil and Wagner 2012, Massa 2013).

The existing macro-level studies face three important limitations. First, while Calì and te Velde (2011) to some extent explored heterogeneity across Aid for Trade categories and across sectors, it remains unclear what strategies within these categories and sectors are most effective. We thus learn about the impact of Aid for Trade in general, but not about the specific impact of Aid for Trade in food value chains through PPPs.

Second, the above studies estimate the effects on trade-related outcomes, but not on progress towards development goals. To what extent Aid for Trade actually contributes to intended development goals has remained an open question.

Third, the above studies rely on country- and sector-level evaluations, they are unable to test micro-economic explanations that may help to understand why the impact of Aid for Trade may vary by category, sector, strategy and intervention.

1.2.2. Specific impact through PPPs in food value chains

This thesis aims to address limitations of current studies by considering the impact of Aid for Trade through PPPs in food value chains specifically. I take a micro-level approach to be able to evaluate its (potential) contribution to increased food safety (SDG 2) and increased farmer incomes (SDG 1), and test micro-economic explanations for the effects that are found.

---

[4] Poulton and Macartney (2012) discuss market failures that affect agricultural markets in general, and therefore also include the lack of enabling infrastructural environment. Since I limit myself to PPPs in food value chains, I omit this category here.

To better understand where PPPs can effectively contribute to development goals, it might be helpful to consider the evidence presented in this thesis through the lens of principal-agent theory. The donor is the principal, and wants the private agent to contribute to development goals. The principal can provide incentives to the agent, but effort is often difficult to contract. The intended outcomes will only be reached if private interests are aligned with development goals. Whenever private interests and development goals are misaligned, intended outcomes are unlikely to be reached (Hart and Holmström 1987, Poulton and Macartney 2012).

1.2.3. Microeconomic explanations

To explore the alignment of private interests with specific development goals, and the potential for PPPs to contribute to these goals, this thesis largely builds on two main strains of microeconomic literature. First, the impact that PPPs can have on food safety depends on how smallholder farmers react to market incentives. While price incentives may exist at downstream levels, smallholders' deliveries are often aggregated quickly after selling to the market. Without individual quality measurement systems, farmers are therefore generally under-incentivized to deliver high-quality food. PPPs can help to extend incentives to smallholder producers. While the literature finds that smallholders react to incentives in general (Casaburi and Macchiavello 2015, Bernard et al. 2019, Burchardi et al. 2019), and price incentives for food safety in particular (Saenger et al. 2013, Bernard et al. 2017, Hoffmann and Jones 2018, Hoffmann, Magnan, et al. 2018), impacts could be limited if quality is assessed by private buyers (Saenger, Torero, and Qaim 2014) or farmers produce mostly for home consumption. This thesis studies to what extent impacts of price incentives for food safety also extend to such situations.

Second, the impact of PPPs on farmer incomes depends on the vertical distribution of rents created by increases in food safety, and thus on the formation of prices and the competitiveness of the market. Whether agricultural markets are competitive is a long-standing and pertinent question. A review by Dillon and Dambro (2017) found no support for widespread rent-extraction by traders, but hardly any evidence exists on rent-extraction by large food processing and exporting companies, which are often involved in PPPs. In fact, widespread rent-extraction by large processing and exporting companies could potentially explain the lack of over-all impacts of aid for productive capacity. This thesis provides some first evidence on rent-extraction by a food processing company.

1.3. Objective and research question

The general research question of this thesis is: "What is the potential for public-private partnerships to support food safety and farmer incomes in food value chains in lower middle-income countries?"

I study the (potential) contribution of PPPs to increased food safety (SDG 2) and increased farmer incomes (SDG 1). More specifically, the five Chapters study:

A. the drivers of food safety technology adoption, both intrinsic drivers (Chapter 2) and market incentives (Chapters 3 and 4),
B. the inclusiveness of PPP interventions (Chapter 6), and
C. the distribution of intervention rents, both vertically between producers and processors (Chapter 5) and horizontally between farmers (Chapter 6).

Food safety can be defined narrowly as "the probability of not contracting a disease as a consequence of consuming a certain food", or more broadly be viewed as "also encompassing nutritional qualities of food" (Grunert 2005). The food safety concept used in Chapter 2 and 3 fits in both of these definitions. Food safety is conceptually different from the broader notion of food quality, which I define for the purpose of this thesis as the "[desirable] physical characteristics built in the product" (Grunert 2005). Chapter 4 uses the concept of food quality, because it is broader, and because it better reflects the final use of improved raw produce. While food safety is more strictly regulated, and consumers may generally assume food to be safe (Hoffmann, Moser, and Herrman 2019), the distinction between food safety and food quality may have strong implications for the marketing to consumers. However, the distinction between food safety and food quality may be less important for the impact of providing market incentives, particularly to market producers, and I therefore review studies on both food safety and food quality to learn about food safety improvement.

The individual Chapters vary in the extent to which they are framed in the narrow debate on PPPs and the wider debate on Aid for Trade in food value chains. Chapters 2, 3 and 4 are not framed specifically in any of these debates, as they contribute to a better general understanding of the process of food safety technology adoption. While these Chapters can inform the design of PPP interventions, their lessons are neither unique to PPPs nor to Aid for Trade, and could be relevant to other actors that wish to improve food safety.

Chapter 5 is framed in the wider debate on Aid for Trade in food value chains, for two reasons. First, for the distribution of rents, it does not necessarily matter whether the quality increase has been facilitated by the PPP or by other actors. Second, while the produced food is most likely not formally exported, it does not necessarily matter for the farmers whether their production is consumed within domestic urban areas, or exported to nearby countries. The findings may therefore generalize to other interventions that support farmers to produce high-quality output in contexts where buyers have market power, and may provide important insights for the wider debate on Aid for Trade in food value chains.

Chapter 6 is framed in the narrow debate on PPPs in food value chains, as it studies the inclusiveness and distributional effects of a PPP intervention, and is therefore relevant specifically to this debate.

1.4. Methodology

This thesis uses microeconomic models and various impact evaluation methods to identify and explain the (potential) contribution of PPPs to support food safety and farmer incomes in food value chains in lower middle-income countries. Theoretical predictions are being tested in empirical settings in the maize value chain in Kenya, one of the less developed lower middle-income countries, and the dairy value chain in Indonesia, one of the more developed lower middle-income countries.

1.4.1. Microeconomic models

Chapter 3, 4 and 5 use microeconomic models to explicitly outline potential mechanisms through which outcomes can be affected, and thereby contribute to the interpretation and explanation of empirical findings. These models make explicit assumptions on preferences of various value chain actors and the costs of producing high-quality and safe food. These assumptions are used to derive predictions on behavior, which are subsequently tested in the empirical settings.

1.4.2. Experiments

Testing these predictions can be complicated. It is often hard or impossible to find credible control groups to evaluate the impact of PPPs. Moreover, PPPs frequently implement packages of interventions (see Chapter 4 for an example), which complicates the attribution of development impacts to specific mechanisms.

To be able to study how PPPs can contribute to development goals, we therefore organized a Randomized Controlled Trial. We designed the experimental market linkage treatments ourselves, so that we could isolate specific mechanisms. And we randomized these experimental treatments across participating villages to create similar treated and control groups. In this way, Chapter 3 is able to causally isolate how specific market linkage mechanisms drive the food safety technology adoption decisions of smallholder farmers.

In the process, we also randomized the baseline survey status of respondents, given some constraints. To learn more about 'soft' drivers of food safety technology adoption, like the attention paid to the food safety issue, Chapter 2 uses the randomized survey status as instrumental variable to causally identify the impact of being surveyed on the adoption of technology. The methodological insights of Chapter 2 are being used in Chapter 3.

1.4.3. Quasi-experiments

Some questions, however, would be impossible or too costly to study in a researcher-led intervention. Chapters 4-6 are therefore based on a PPP-led intervention that intended to increase the quality of milk and to contribute to increased profitability of dairy farms.

As the intervention under scope was not randomized, Chapter 6 is able to study the inclusiveness of this intervention and explore how pre-existing inequality between farmers is affected.

Chapter 4 specifically studies the impact of the intervention on milk quality. Since the intervention comprises of a package of sub-interventions, attribution to specific mechanisms is complicated. However, since some sub-interventions could only affect specific quality dimensions, I am able to use multiple quality dimensions to learn about the likely mechanisms that have driven the effects.

To create similar intervention groups for the analyses in Chapters 4 and 6, I employed a Coarsened Exact Matching procedure (Iacus, King, and Porro 2012). The panel nature of the administrative data used in these papers allowed me to show trends over time, which provide confidence in the causal interpretation of results. Due to the small number of clusters, p-values are bootstrapped (Cameron, Gelbach, and Miller 2008).

In the context of potential monopsonistic power, Chapter 5 subsequently argues that the effect on prices can only be evaluated at the aggregate level, and therefore uses a Difference-in-

Differences design to study the effect on prices. Being unique in its identification strategy, Chapter 5 draws some important methodological lessons for the evaluation of impacts on prices, which are incorporated in the empirical design of this Chapter, and extensively discussed in the Discussion of Chapter 5 and the General Discussion of this thesis.

1.5. Outline of the thesis

The remainder of this thesis is structured as follows. Chapters 2 and 3 are based on a project in the maize value chain in Kenya, where a team of researchers studied how subsistence farmers can be supported to improve food safety through adoption of a new technology.

Chapter 2 studies the impact of being surveyed on the adoption of food safety technology, and finds that it has a large effect on adoption, at both the intensive margin (the intensity of adoption) and the extensive margin (whether one adopts or not). The Chapter suggests that the impact is caused by focusing the farmers' attention to the issue of food safety. It also shows that experimental parameters of interest can be affected by surveying farmers, and therefore informs the empirical strategy of Chapter 3.

Chapter 3 is based on joint work with Vivian Hoffmann, Sarah Kariuki and Janneke Pieters, and studies the effect of providing market premiums for safe maize on the adoption of a food safety technology among subsistence farmers in Kenya. We model the uptake of food safety technology under the production risk. We randomized the provision of the market premium across villages, and find in line with model predictions that food safety technology adoption increased at the intensive, but not at the extensive margin.

Chapters 4-6 are based on a PPP's intervention that aimed to increase food quality and increase farm profitability in the dairy value chain in Indonesia. The intervention (i) trained farmers, (ii) upgraded Milk Collection Point facilities, and (iii) introduced an individual quality testing and price incentive system, which replaced a group-level system.

Chapter 4 studies how the intervention affects quality, and finds a positive impact at multiple dimensions of milk quality, which can partly be attributed to changes in the incentive system.

Chapter 5 argues that market power of the buyer can cause the economic benefits of the intervention to be extracted. I model the distribution of intervention rents between the processors and producers, and use a panel of transactions between the dairy cooperative and two processing companies to empirically evaluate the effect on prices. The Chapter finds that prices are temporarily higher, but that intervention rents are extracted before the cooperative earned back its portion of the investment in the intervention.

Chapter 6 is based on joint work with Jos Bijman, Erwin Bulte and Marlene Roefs, and explores the inclusiveness of this intervention. The Chapter finds that more developed farmers with more cows and larger Milk Collection Points are more likely to be reached by the intervention, and shows that income-inequality between farmers therefore increases.

Finally, the General Discussion synthesizes the available evidence, provides recommendations for public-private interventions in food value chains, and discusses considerations for further research.

2. The Impact of Being Surveyed on the Adoption of Agricultural Technology

Mark Treurniet

2

This paper uses exogenous variation in the probability of being surveyed at baseline to estimate the impact of being surveyed on subsistence farmers' take-up of a new agricultural technology that improves food safety. I find large and statistically significant impacts of being surveyed, and also find that an experimental treatment effect disappears for surveyed farmers. My results have strong implications for our understanding of the process of technology adoption, for the external validity of adoption results measured in surveyed populations, and for research ethics.

2.1. Introduction

While the adoption of new technologies is an important driver of growth in output and quality of agricultural production, adoption has remained low in many developing countries (Foster and Rosenzweig 2010). Many studies therefore explore various market inefficiencies that may constrain agricultural technology adoption (Jack 2013). Often, such studies involve household surveys. The effects of such surveys on later technology adoption are understudied.

Being surveyed may affect take-up of agricultural technologies in several, possibly related ways. First, surveys may focus scarce cognitive capacity and executive control to food safety risks. The availability of such "bandwidth" may play an important role in technology adoption (Schilbach, Schofield, and Mullainathan 2016). Second, question-behavior effects arise if answering questions on predictions or intentions on specific behavior affects behavior (Rodrigues et al. 2015). Third, being surveyed may make farmers more aware that their behavior is observed as part of a study. Hawthorne effects arise if this increased awareness changes adoption behavior (McCarney et al. 2007). Fourth, experimenter demand effects emerge if surveys change the respondent's perception of what the experimenter regards as "appropriate" behavior and this affects take-up decisions (De Quidt, Haushofer, and Roth 2018).

While previous studies find that questions-behavior effects and experimenter demand effects are generally limited (Rodrigues et al. 2015, De Quidt, Haushofer, and Roth 2018), Schilbach, Schofield, and Mullainathan (2016) argue that the availability of bandwidth may be especially important for technology adoption processes. This paper studies the impact of being surveyed on technology adoption among subsistence farmers, and explores how the interaction between being surveyed and experimental treatments can affect the estimation of parameters of interest.

In a previous study, Zwane et al. (2011) report five different field experiments on the impact of being surveyed about health and/or household finances. In three health experiments, they find that being surveyed increases the use of water treatment products and take-up of medical insurance. In two microfinance experiments, they do not find an effect of being surveyed on borrowing behavior. The authors speculate that these results can be explained by what Schilbach, Schofield and Mullainathan later defined as bandwidth.

The objective of this paper is to explore the external validity of the results of Zwane et al. (2011), and study the impact of being surveyed on the take-up of a new agricultural technology that improves food safety. As part of a larger baseline survey, questions were asked on a specific food safety issue and experience with prevention measure, but no questions were asked on predictions or intentions to adopt a newly available technology. I use randomized variation in the probability of being surveyed at baseline to find large and statistically significant impacts of being surveyed on both the extensive and intensive margin of adoption, as recorded for both surveyed and non-surveyed farmers during sales. Moreover, I find that conducting baseline surveys affects the estimation of experimental parameters of interest: The experimental treatment effect in a related experiment disappears for surveyed respondents.

The remainder of this paper is structured as follows: The next section discusses the context in which the impact of being surveyed is studied, as well as the randomization procedure and available data. Subsequently, I discuss empirical strategy and results, first for the impact of being surveyed, and then for the interaction with related parameters. The last section discusses implications.

2.2. Context, experiment and data

Aflatoxin is a fungal toxin that can grow in maize and groundnuts. High levels of exposure to aflatoxin may cause cancer, liver damage and death (Strosnider et al. 2006). The field experiment described in Chapter 3 studies constraints to subsistence farmers' adoption of Aflasafe™, a biocontrol product that decreases aflatoxin contamination in maize. Aflasafe™ was recently introduced in Kenya, the country in which half of 152 pre-existing producer groups were offered a guaranteed bonus for safe maize. This treatment was randomized across groups.

Prior to project trainings and Aflasafe™ sales, a baseline survey was conducted between September 14 and October 16, 2017. The baseline survey included questions on household demographics, household and livestock assets, land ownership and use, agricultural input use, maize harvest, post-harvest handling of maize, maize marketing, maize sale, expectations for the coming season, income sources, risk and insurance, group membership, as well as some questions on knowledge of aflatoxin and experience with aflatoxin prevention measures.

For budgetary reasons, not all farmers were selected for baseline surveys. The sampling procedure created random variation in the probability of being surveyed. First, six primary members per group were randomly selected for interviews. These farmers were called 2-3 days before the day of the surveys. Subsequently, six members were randomly selected to replace the initially selected members if interviews could not be completed. If less than six surveys could be completed among primary respondents on the day of the surveys, enumerators started calling replacement respondents from the top of the replacement list.

While the sampling procedure created random variation in the probability of being surveyed, it also prioritized some members over others, for which I will control in the analysis. During a group census that had taken place between April 29 and August 25, 2017, a list was made of all group members. Baseline survey respondents were first selected from members that had been present during the group census meeting. When this pool was exhausted, respondents were subsequently selected from members that had not been present during this meeting. Only the first 20 present members listed within each group were included in the randomization, and the remaining members were not considered. If the present pool was exhausted before the required twelve members were selected, then only the first 20 non-present members listed were included in the randomization of selection of the remaining respondents, and the remaining non-present members were not considered. In the analysis for this paper, I include only those individuals that were included in the randomization,[5] and the econometric specification controls for the prioritization of present members over non-present members.

Figure 2.1 shows the proportion of being surveyed by survey selection status. Being selected in our primary sample increases the probability of being surveyed at baseline by 0.728 as compared to not being selected for surveying. As expected, being selected as replacement still increases the probability of being surveyed at baseline, but the effect decreases with the rank ($p = 0.000$). Conditionally on being called, the probability of being surveyed is higher for primary respondents than for replacement respondents ($p = 0.000$), but does not differ among replacement respondents with different rank ($p = 0.4543$).

During the sale of Aflasafe™, which took place between November 10 and December 18, 2017, the identities of buyers were again recorded. I am therefore able to match the

[5] For one group, the group census list was lost and re-taken later, making it impossible to retrieve which farmers were considered for surveys and which farmers were not. I therefore exclude this group from this analysis. Since I use individual level variation in this analysis, excluding this one group does not affect the internal validity of the estimated impact of being surveyed.

administrative sales data to the randomized sample selection and actual survey completion statuses.

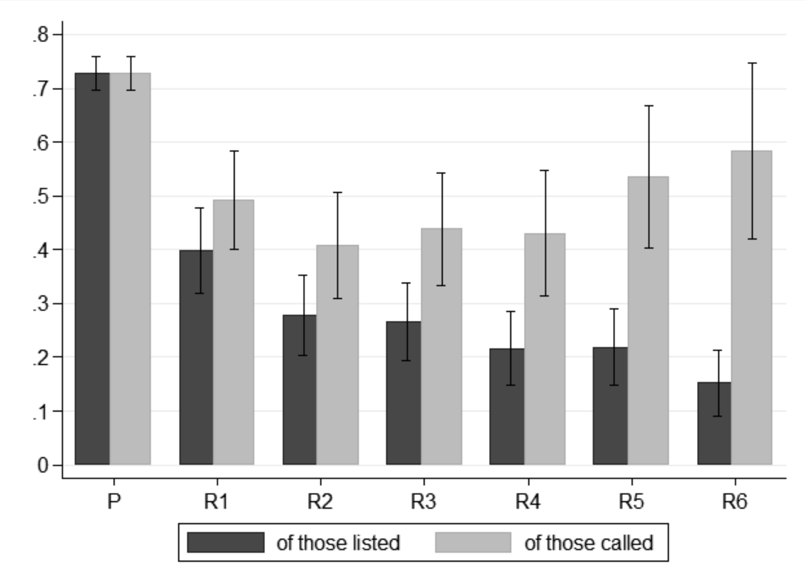

Figure 2.1. Proportion surveyed

2.3. Impact of being surveyed

2.3.1. Empirical strategy

To study the impact of being surveyed on the adoption of Aflasafe™, I estimated:

$$Adoption_{ig} = \beta Surveyed_{ig} + \gamma_g + \delta_g Present_i + \varepsilon_{ig}, \qquad (2.1)$$

where $Adoption_{ig}$ is the actual adoption of Aflasafe™ of individual i from group g, which I will analyze at both the extensive margin (whether the individual purchased Aflasafe™ or not) and the intensive margin (the quantity of Aflasafe™ purchased in kg), $Surveyed_{ig}$ is a dummy indicated whether the individual is surveyed at baseline, and ε_{ig} is an error term. As survey status was randomized within groups and respondents were first selected from present members and subsequently from non-present members, I included separate group fixed effects for

2

present and non-present members $\gamma_g + \delta_g Present_{ig}$ to control for randomization strata.[6] Note that the experimental treatments described in Chapter 3 were randomized at the group level, and are therefore captured in γ_g.

As there is partial compliance to the randomized survey status, actual survey completion is likely to be endogenous. I therefore instrument $Surveyed_{ig}$ by a dummy for selection into the primary sample and six dummies for selection as replacement, while expecting the probability of being surveyed to be highest for the primary sample and decreasing in the rank for replacements. More formally, I estimated:

$$Surveyed_{ig} = \lambda_0 Primary_i + \sum_{r=1}^{6} \lambda_r \, Replacement_i + \rho_g + \tau_g Present_{ig} + \omega_{ig}, \qquad (2.2)$$

where $Primary_i$ is a dummy for selection into our primary sample, $Replacement_i$ is a dummy for selection as r^{th} replacement and ω_{ig} is an error term.

Results were obtained using two-stage least squares (2SLS). Since groups were sampled from a larger population and the treatment effect might be heterogeneous, standard errors were clustered at the group level (Abadie et al. 2017).

2.3.2. Results

Table 2.1 reports our results on the impact of being surveyed. I find large and statistically significant impacts of being surveyed. Adoption increased by 52% at the extensive margin and by 74% at the intensive margin as compared to the control group.

[6] Additionally controlling for gender does not affect my results.

Table 2.1. Impact of Being Surveyed

	Outcome variables	
	(1)	(2)
	Adoption	Intensity (kg)
Surveyed	0.097***	0.370***
	(0.035)	(0.111)
Groups	151	151
Observations	2563	2563
Mean of non-surveyed	0.185	0.501

2SLS results, with randomized respondent selection instrumenting for survey completion, and controls for randomization strata
Standard errors clusted at group level in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Mean of non-surveyed calculated as sample mean -/- proportion surveyed * estimated impact of being surveyed

2.3.3. Robustness

During project trainings, information was given on Aflasafe™, the rainfall index insurance sold together with Aflasafe™ and, when relevant, the guaranteed bonus for safe maize. Since these trainings were often organized at the same location and day as baseline surveys, being surveyed could be correlated with presence at this training. If presence at the training subsequently increased the adoption of Aflasafe™, then our instrumental variables are invalid to estimate the effect of being surveyed on the adoption of Aflasafe™. More specifically, our 2SLS estimate would capture:

$$E[\beta] = \frac{p_s\beta_s + (p_{t,1} - p_{t,0})\beta_t}{p_s} = \beta_s + \frac{p_{t,1} - p_{t,0}}{p_s}\beta_t, \qquad (2.3)$$

where p_s is the probability of being surveyed, β_s is the impact of being surveyed, $p_{t,1}$ is the probability of presence at the training given that the farmer is selected to participate in the survey, $p_{t,0}$ is the probability of presence at the training given that the farmer is not selected to participate in the survey and β_t is the impact of presence at the training; and I assume that $p_{t,1} \geq p_{t,0} > 0$ and $\beta_t \geq 0$. Clearly, the 2SLS estimate is biased if $p_{t,1} > p_{t,0}$ and $\beta_t > 0$.

As primary respondents were called 2-3 days before the day of the surveys and necessary replacement respondents were only called at the day of the surveys, primary respondents were more likely to be surveyed than replacement respondents (see Figure 2.1). In the extreme case that $p_{t,1}$ and p_s are perfectly correlated, the relative effect $(p_{t,1} - p_{t,0})/p_s = (p_s - p_{t,0})/p_s$ is larger for farmers on the primary list than for farmers on the replacement list. If presence at the training subsequently increased the adoption of Aflasafe^{TM}, I would then expect the 2SLS estimate to be more positively biased for primary respondents. As a robustness check, I therefore estimated our 2SLS specification including the instrumented interaction between being surveyed and primary respondent status, so that I can compare 2SLS estimates across primary and replacement respondents.

Table 2.2 shows the results of this estimation. The interaction effect between being surveyed and being a primary respondent is not significantly different from zero and the point estimate is negative. I thus find no indication for the 2SLS estimate to be biased.

Table 2.2. Robustness of Impact of Being Surveyed

	Outcome variables	
	(1)	(2)
	Adoption	Intensity (kg)
Surveyed	0.168**	0.626**
	(0.077)	(0.245)
Surveyed*Primary	-0.061	-0.222
	(0.062)	(0.204)
Groups	151	151
Observations	2563	2563
Mean of non-surveyed	0.177	0.470

2SLS results, with randomized respondent selection instrumenting for survey completion, and controls for randomization strata
Standard errors clusted at group level in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Mean of non-surveyed calculated as sample mean -/- proportion surveyed * estimated impact of being surveyed -/- proportion primary and surveyed * estimated interaction effect

An alternative explanation for the absence of an instrumented interaction effect with primary respondent status is that surveyed primary respondents structurally differ from surveyed replacement respondents, because they were called 2-3 days before the day of the surveys. Table 2.3 therefore reports differences across primary and replacement respondents on a selection of survey variables that was also used in Chapter 3. I do not find structural differences between primary and replacement respondents, and thus do not find empirical support for this alternative explanation.

Table 2.3. Balance Across Primary and Replacement Respondents

	Primary			Replacement			Diff	
	N	Mean	SD	N	Mean	SD	P[1]	P[2]
Age of the farmer (completed years)	652	50.1	13.8	220	50.4	14.3	0.793	0.086
Years of education completed by head	652	7.13	4.04	220	7.16	3.94	0.930	0.971
Relationship with the head	652	0.586	0.493	220	0.582	0.494	0.916	0.300
Asset index	652	5.61	2.33	220	5.88	2.29	0.142	0.448
Total land under maize main season previous year (acre)	652	1.43	1.20	220	1.56	1.32	0.191	0.629
Maize harvest main season previous year (kg)	652	433	777	220	474	715	0.475	0.413
Maize marketing: whether sold any maize last season	652	0.465	0.499	220	0.495	0.501	0.431	0.260
Total expenditures on agr. inputs & labour main season previous year (KES)	652	10832	11894	220	11242	11087	0.642	0.684
Propensity for social learning dummy	652	0.495	0.500	220	0.418	0.494	0.046	0.052
Aflatoxin knowledge index	652	-0.030	0.812	220	0.030	0.745	0.310	0.199
Knowledge and experience with insurance	652	1.33	0.82	220	1.30	0.83	0.679	0.409
Individual trust index	652	0.003	0.548	220	0.008	0.563	0.904	0.931
Punishment index	652	-0.041	1.007	220	0.098	0.968	0.069	0.799
Quantitative risk aversion score	652	4.43	2.84	220	4.17	2.91	0.247	0.414

[1] P-value based on difference between primary and replacement and robust standard errors

[2] P-value based on difference between primary and replacement after controlling for randomization stata, and robust standard errors

2

2.4. Interaction with experimental treatment

2.4.1. Empirical strategy

To explore the interaction of being surveyed and experimental parameter estimates, I study how being surveyed affects the impacts of the market linkage treatment studied in Chapter 3.[7] I therefore estimated:

$$Adoption_{igv} = \alpha Treated_v \cdot Surveyed_{igv} + \beta Surveyed_{igv} + \gamma_g + \delta_g Present_i + \varepsilon_{igv}, \quad (2.4)$$

where $Treated_v$ is a dummy representing the market linkage treatment, which was randomized at the village level v, and ε_{igv} is an error term. Note that the market linkage treatment was randomized at the group level, and is therefore again captured by γ_g. Together with $\delta_g Present_i$, γ_g is necessary to control for randomization strata.

To deal with endogeneity in survey completion, I instrumented $Surveyed_{igv}$ and $Treated_v \cdot Surveyed_{igv}$ by the primary respondent status, the six replacement respondent statuses, and their interactions with the market linkage treatment.

Results were obtained using 2SLS and standard errors were clustered at the village level.

2.4.2. Results

Table 2.4 reports the results on the interaction between the market linkage treatment and being surveyed. I am not able to establish a significant interaction effect due to large standard errors caused by comparing the IV estimate across a treatment randomized at the village level. In fact, given these standard errors, it would even not be possible to establish a significant interaction effect if the market linkage treatment would not at all have affected surveyed individuals and it would have affected non-surveyed individuals so much that it can completely explain the overall market linkage effect on the intensity of adoption, as found by Chapter 3.

[7] Appendix Table A2.1 shows that I find a similar pattern of treatment effects as Chapter 3 if I re-estimate the impact of the market linkage treatment for the sample used in this paper, which excludes members not included in the randomized selection of baseline respondents.

Table 2.4. Interaction with Market Linkage Treatment

	Outcome variables	
	(1)	(2)
	Adoption	Intensity (kg)
Market linkage*Surveyed	-0.069	-0.225
	(0.072)	(0.209)
Surveyed	0.132**	0.484***
	(0.054)	(0.134)
Villages	123	123
Observations	2563	2563

2SLS results, with randomized respondent selection instrumenting for survey completion, and controls for randomization strata
Standard errors clusted at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01

However, the imprecisely estimated interaction effect is so large that a market linkage effect would not be found if the analysis focuses on surveyed farmers only. This pattern is in line with the results in Zwane et al. (2011), where an experimental treatment effect on technology adoption disappears for surveyed respondents. Although I cannot formally attribute causality, the interaction results suggest that being surveyed biases the market linkage treatment effect.

2.5. Discussion

This paper finds large and significant impacts of being surveyed on subsistence farmers' adoption of a new agricultural technology that improves food safety. While this paper cannot isolate the different channels, I speculate in line with previous work (Zwane et al. 2011, Rodrigues et al. 2015, De Quidt, Haushofer, and Roth 2018) that the bandwidth channel is an important driver of the impact of being surveyed.

Following Zwane et al. (2011), my findings have crucial implications for the study of technology adoption. Substantively, as already suggested by Schilbach, Schofield, and Mullainathan (2016), bandwidth seems to be an important factor for the study of technology adoption. If insufficient bandwidth (cognitive capacity and executive control) is dedicated to a specific issue, people might ignore new technologies that address this issue. Further studying the formation and deployment of bandwidth might improve our understanding of technology adoption.

Methodologically, studies that rely exclusively on surveyed samples are likely to provide adoption estimates that are higher than they would be in non-surveyed populations. Moreover, if survey effects and other treatment effects are not additively separable, estimates of treatment effect for surveyed samples may not be valid for external populations. Biases may especially arise in situations where available bandwidth is an important driver of technology adoption, and financial costs and benefits play a smaller role. If controlling for baseline covariates is still preferred in such settings, and baseline covariates are likely to be correlated within groups, one can consider to survey a subset of the members of each group and use group level means of baseline covariates in the analysis of adoption among the remaining group members (Chapter 3).

Ethically, while the efficacy of new technologies may not yet have been tested outside controlled agronomic experiments, researchers directly affect the adoption of technologies. If this investment does not pay-off, welfare is affected negatively. Moreover, bandwidth is a scarce resource and using some bandwidth for one task may leave less for other tasks (Mani et al. 2013). This might lead to worse over-all outcomes. Further research should shed more light on such undesirable side-effects.

Appendix

Table A2.1. Impact of Market Linkage Treatment in Sample Used in This Paper

	Outcome variables			
	(1)	(2)	(3)	(4)
	Adoption	Adoption	Intensity (kg)	Intensity (kg)
Market linkage	0.022	0.020	0.217**	0.178*
	(0.029)	(0.030)	(0.099)	(0.093)
Baseline controls	No	Yes	No	Yes
Villages	123	123	123	123
Observations	2563	2563	2563	2563
Mean of no market linkage	0.207	0.207	0.520	0.520

Standard errors clusted at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01

2

3. Safety for My Food – and Maybe Yours Too: Upside Risk and the Impact of Premium Market Access on Subsistence Producer Demand for a Food Safety Technology

Vivian Hoffmann, Sarah Kariuki, Janneke Pieters and Mark Treurniet

Premium prices conditional on food safety attributes are often proposed as a way to increase food safety for marketed produce. This paper studies how a quality premium affects the adoption of food safety technology among subsistence farmers with a stochastic harvest. We present a simple model showing that a modest quality premium that is too low to affect adoption on the extensive margin can harness upside risk by providing subsistence farmers a high-value market for their excess high-quality output. We test our model's predictions in a randomized field experiment among maize farmers in Kenya, and find in line with our model predictions that a modest quality premium for food safety increases subsistence farmers' adoption of an aflatoxin-reducing technology at the intensive margin, but not at the extensive margin.

3.1. Introduction

Foodborne pathogens and toxins exact a significant health toll in developing countries (WHO 2015), particularly among the poorest (Leroy, Wang, and Jones 2015). Further, food safety is an increasingly important precondition for access to high value markets (Ashraf, Giné, and Karlan 2009, Van Beuningen and Knorringa 2009). Improving the safety of food produced for subsistence typically relies on education and subsidies. Regulatory or voluntary standards, the latter combined with price premiums or other market advantages, are employed in the case of marketed food. In this paper, we examine how a market-based instrument – a price premium for food safety – affects the safety of food produced for home consumption in the context of production risk. We show that when production quantity is stochastic, access to a premium market for safe food reduces expected exposure to a common food safety hazard among subsistence producers.

We consider the case of aflatoxin, a common mycotoxin, in maize produced by Kenyan smallholder farmers. Dietary aflatoxin exposure causes cancer and liver damage (Strosnider et al. 2006), and evidence is emerging that the toxin plays a role in childhood stunting (Gong et al. 2004, Turner et al. 2007, Hoffmann, Jones, and Leroy 2018). At high levels of exposure, aflatoxin can cause jaundice, permanent liver damage, and death. Dozens of cases of acute aflatoxin poisoning have been linked to the consumption of maize produced and stored by households in eastern Kenya, the setting of our study (Daniel et al. 2011).

Widespread contamination of maize and groundnut with the toxin limits African nations' export opportunities (Munasib and Roy 2011). Increasingly, domestic food processors vulnerable to reputation effects and international humanitarian organizations required to adhere to global standards in the food they distribute avoid sourcing from aflatoxin-affected regions (Hoffmann and Moser 2017). Some processors in Africa go as far as to import aflatoxin-susceptible ingredients from the Americas at considerable cost (personal communication, Carly Edwards, Project Peanut Butter, March 28, 2018).

Effective technologies to reduce aflatoxin contamination are available, but face several barriers to adoption. First, food safety is a hidden trait, and its observation requires specialized tests that are costly relative to the value of farm produce. In the case of aflatoxin, a single test costs on the order of US $10 – up to half the cost of the typical value of maize grains sold by smallholder farmers in at a time.[8] This information problem is exacerbated by the fact that Kenyan maize supply chains often include multiple intermediaries.

However, as consumer awareness and regulatory capacity to address food safety increase, incentives are growing for Kenyan maize processors to secure safer inputs by establishing direct procurement relationships with farmer groups. Several studies have shown that market incentives affect production and marketing decisions in general (Casaburi and Macchiavello 2015, Bernard et al. 2019, Burchardi et al. 2019), and for food quality and safety specifically (Saenger et al. 2013, Bernard et al. 2017, Hoffmann and Jones 2018, Hoffmann, Magnan, et al. 2018, Chapter 4). However, previous studies of the adoption of food safety and quality technologies have either focused on marketed produce (Saenger et al. 2013, Bernard et al. 2017, Chapter 4), considered technologies for which adoption is a binary decision (Hoffmann, Magnan, et al. 2018), or offered an unrealistically large price premium on a fixed amount of produce, enabling analysis of adoption only on the extensive margin (Hoffmann and Jones 2018). In contrast, we study how a modest market premium affects the adoption of a divisible food safety technology among smallholder farmers.

Another barrier to adoption is production risk. Like many food safety technologies, one of the most effective tools for aflatoxin prevention, the biocontrol product Aflasafe™, is applied during production, before the outcome of this process is observed. In a stochastic agricultural production function, any costly input increases the variability of farm profit (Just and Pope 1978). For low-income populations engaged in rainfed agriculture, who lack access to financial smoothing instruments, this implies increased consumption risk and thus constitutes an important impediment to technology adoption (Rosenzweig and Binswanger 1993, Dercon and Christiaensen 2011, Emerick et al. 2016).

While the literature on how downside risk affects adoption of agricultural technologies is vast, upside production risk has received far less attention. In our setting, upside production risk constitutes a potential driver of adoption: when weather conditions are favorable, farmers harvest more safe grain for a given cultivated area to which a food safety input is applied. If production exceeds household subsistence needs, and no market reward for quality exists, a

[8] The median seasonal sale volume in the counties of Kenya where this research was conducted is 100 kg (Hoffmann and Jones 2018). Beyond the cost of testing supplies, tests should be executed by an experienced technician and compared regularly against results using a reference material to obtain reliable results.

portion of the value of the food safety investment is lost. Below, we present a simple model showing that a modest quality premium that is too low to affect adoption on the extensive margin can harness this upside risk by providing subsistence farmers a high-value market for their excess high-quality output.

We subsequently test the model's predictions by studying the impact of a market premium for food safety on subsistence farmers' adoption of an aflatoxin-reducing technology (Aflasafe KE01^{TM}) through a randomized trial in which farmers in one of the most aflatoxin-affected regions in the world were given the opportunity to purchase Aflasafe^{TM} under experimentally varied market conditions. Half of 152 pre-existing producer groups were assigned to a market linkage treatment and offered a premium price for the maize they aggregated if it conformed to the East African aflatoxin standard.[9]

We find that the price premium, which was set to a modest 5% of the value of maize, did not affect the extensive margin of adoption, suggesting that farmers who purchased the product used it first on maize produced for their own consumption. We do, however, see a strong positive impact of the premium on the intensive margin of adoption. Farmers who were offered the food safety premium purchased nearly twice as much Aflasafe^{TM} as those not given this opportunity.

We begin by presenting a simple model of the food safety investment decision faced by a subsistence farmer in the context of production risk in Section 3.2. In Section 3.3, we describe the market context, technology offered, and study population. Section 3.4 describes the study design and data, and Section 3.5 outlines the empirical strategy. Results are presented in Section 3.6, and Section 3.7 offers concluding remarks.

---

[9] The market linkage treatment was cross-cut with a bundled insurance treatment, in which AflasafeTM could only be purchased together with an actuarially fair rainfall index insurance product designed to insure against maize losses due to unfavorable weather conditions during the growing period. Farmers not assigned to the bundled insurance treatment who purchased AflasafeTM were able to purchase the same insurance separately. The bundled insurance treatment is described in Hoffmann, Kariuki et al. (2018). As farmers not assigned to the bundled insurance treatment also had the option of buying insurance, and 75% did so, bundling insurance had no impact on adoption.

3.2. Model

In this section we formally model how a modest price premium that is insufficient to induce farmers to adopt a food safety technology at the extensive margin can increase adoption among subsistence farmers on the intensive margin. Intuitively, the price premium increases the value of safe maize produced in excess of home consumption needs. This causes a marginal increase in the expected benefit of investing in safe food, leading the farmer to invest more. We first define the farmer's utility as a function of food safety investments, and then derive conditions for the optimum investment with and without a price premium.

3.2.1. Set-up

Assume that farmers maximize their utility:

$$\max E[U] = E[V] + E[R] - C\,, \qquad (3.1)$$

Where V is the total value of home consumption, R is total revenue of produce delivered to the market, and C is the total cost of investment in food safety.

Let $I \in [0,1]$ denote the proportion of land to which the food safety technology is applied. Let $c > 0$ denote the cost of investment to cover the entire cultivated area. Then, the cost of investment is:

$$C = cI, \qquad (3.2)$$

So that the marginal cost of investment equals c. We assume (i) that investment in food safety directly results in safe produce $s = I$, where s is the proportion of food produced that is safe, (ii) that the total harvest amount is stochastic and uniformly distributed $q \sim U(q_L, q_H)$, so that the mean harvest equals $\mu_q = (q_L + q_H)/2$, (iii) that home consumption varies with harvest, but never exceeds a fixed amount $q_{home} = \min\{q, \widetilde{q_{home}}\}$, (iv) that $\widetilde{q_{home}} \leq q_H$ to generate upside risk, and (v) that the remainder $q_{market} = q - q_{home}$ is sold.

Safe food produced will either be consumed at home or delivered to the market:

$$s_{home}q_{home} + s_{market}q_{market} = Iq, \qquad (3.3)$$

where $s_{home} \in [0,1]$ is the proportion of produce consumed by the household that is safe, and $s_{market} \in [0,1]$ is the proportion of produce delivered to the market that is safe. The farmer first chooses the level of investment I, and then chooses both s_{home} and s_{market} after the realisation of q.

The total value of home consumption equals:

$$V = (\alpha + \beta s_{home})q_{home}, \qquad (3.4)$$

where α is the value of consuming food of the quality produced by the farmer in the absence of any food safety investment, and β is the value premium for consuming safe food.

3

Farmers vary in the additional value they derive from consumption of safe food:

$$\beta \in \{\beta_L, \beta_H\}, \qquad (3.5)$$

where, to make the model interesting, we assume that:

$$\beta_L < \frac{c}{\mu_q} < \beta_H < \tilde{\beta} \equiv \frac{c}{\left.\frac{\partial E[s_{home}q_{home}]}{\partial I}\right|_{I=1}}, \qquad (3.6)$$

The first inequality ensures that the low type will not adopt the technology, and the second inequality ensures that the high type will adopt in absence of a premium price. The third inequality ensures that the high type's motivation to produce safe food for home consumption is by itself insufficient for full adoption, so that some room is left for market incentives to increase adoption at the intensive margin. One can later easily see that $\tilde{\beta} = \infty$ if $q_L \geq \widetilde{q_{home}}$.

The total revenue of produce delivered to the market equals:

$$R = (\gamma + \delta s_{market})q_{market}, \qquad (3.7)$$

where γ is the standard commodity price, and δ is the price premium for safe produce delivered to the market.

The farmers' utility maximization problem then becomes:

$$\begin{aligned} \max_{I, s_{home}, s_{market}} \quad & E[U] = \alpha E[q_{home}] + \beta E[s_{home}q_{home}] \\ & + \gamma E[q_{market}] + \delta E[s_{market}q_{market}] - cI, \qquad (3.8) \end{aligned}$$

subject to (3.3), and $s_{home}, s_{market}, I \in [0,1]$.

3.2.2. Solution

First consider $\delta = 0$, so that the only motivation to make a costly investment in food safety comes from β. The farmer will select the safest produce for home consumption and deliver the remainder to the market, so that:

$$s_{home}q_{home} = \min\{Iq, q_{home}\} = \min\{Iq, \widetilde{q_{home}}\} = \begin{cases} Iq & if\ q \leq \dfrac{\widetilde{q_{home}}}{I} \\ \widetilde{q_{home}} & if\ q > \dfrac{\widetilde{q_{home}}}{I} \end{cases}, \qquad (3.9)$$

which in expectation equals:

$$E[s_{home}q_{home}] = \begin{cases} I\mu_q & if\ I \leq \dfrac{\widetilde{q_{home}}}{q_H} \\ f(I) & if\ \dfrac{\widetilde{q_{home}}}{q_H} < I \leq \dfrac{\widetilde{q_{home}}}{q_L} \\ \widetilde{q_{home}} & if\ I > \dfrac{\widetilde{q_{home}}}{q_L} \end{cases}, \qquad (3.10)$$

3

where:

$$f(I) = \int_{q_L}^{\widetilde{q_{home}}/I} Iq \frac{1}{q_H - q_L} dq + \int_{\widetilde{q_{home}}/I}^{q_H} \widetilde{q_{home}} \frac{1}{q_H - q_L} dq\,, \qquad (3.11)$$

Equation (3.10) is differentiable and monotonically increasing in I, and its first derivative with respect to I equals:

$$\frac{\partial E[s_{home}q_{home}]}{\partial I} = \begin{cases} \mu_q & if\ I \leq \dfrac{\widetilde{q_{home}}}{q_H} \\ \dfrac{1}{2}\dfrac{(\widetilde{q_{home}}/I - q_L)(\widetilde{q_{home}}/I + q_L)}{q_H - q_L} & if\ \dfrac{\widetilde{q_{home}}}{q_H} < I \leq \dfrac{\widetilde{q_{home}}}{q_L} \\ 0 & if\ I > \dfrac{\widetilde{q_{home}}}{q_L} \end{cases}, \qquad (3.12)$$

which is continuous and monotonically decreasing in I.

Intuitively, equation (10) says that when investment in food safety is so low that there is not sufficient safe produce to satisfy home consumption needs even in the case of the highest possible harvest, all safe food is consumed by the household. Beyond this level of investment, the expected quantity of safe home consumption is increasing with investment in food safety, but at a decreasing rate, since the greater the share of land to which the technology is applied, the higher the chance of producing more than is needed for household consumption. Eventually, investment in the food safety technology reaches a point at which there will always be sufficient safe produce for home consumption even when the minimum production value, q_L, is realized. However, this last point will never be reached if $\widetilde{q_{home}} > q_L$.

For farmers who place a low value, β_L, on the safety of home consumption, we have that:

$$\beta_L \frac{\partial E[s_{home} q_{home}]}{\partial I} \leq \beta_L \mu_q < c, \qquad (3.13)$$

which implies that the marginal benefits of investment in food safety are strictly smaller than the marginal costs of investment. These farmers will not invest in food safety.

3

For farmers who place a high value, β_H, on the safety of home consumption, we have $\beta_H \mu_q > c > 0$ and:

$$\beta \frac{\partial E[s_{home} q_{home}]}{\partial I}\bigg|_{I=1} < c, \qquad (3.14)$$

so that the optimal investment I^* is uniquely defined by:

$$\beta_H \frac{\partial E[s_{home} q_{home}]}{\partial I}\bigg|_{I=I^*} = c, \qquad (3.15)$$

Intuitively, farmers with a high value of safe home consumption invest in food safety until the probability that additional safe harvest will be consumed at home becomes too low to justify the cost of investment.

Now consider $c/\mu_q > \delta > 0$. As the price premium is insufficient to produce safe food solely for marketing purposes in expectation, it will not induce farmers who place a low value on the safety of home consumption β_L to start investing. This implies that the price premium has no effect at the extensive margin of investment.

Since $\beta_H > c/\mu_q > \delta$, farmers whose valuation of safety of home consumption is β_H will still select the safest produce for home consumption and deliver the remainder to the market, so that equations (3.9) to (3.12) still hold, and:

$$s_{market}q_{market} = \max\{0, Iq - \widetilde{q_{home}}\} = \begin{cases} 0 & if\ q \leq \frac{\widetilde{q_{home}}}{I} \\ Iq - \widetilde{q_{home}} & if\ q > \frac{\widetilde{q_{home}}}{I} \end{cases}, \qquad (3.16)$$

which in expectation equals:

$$E[s_{market}q_{market}] = \begin{cases} 0 & if\ I \leq \frac{\widetilde{q_{home}}}{q_H} \\ g(I) & if\ \frac{\widetilde{q_{home}}}{q_H} < I \leq \frac{\widetilde{q_{home}}}{q_L}, \\ I\mu_q - \widetilde{q_{home}} & if\ I > \frac{\widetilde{q_{home}}}{q_L} \end{cases} \qquad (3.17)$$

where:

$$g(I) = \int_{\widetilde{q_{home}}/I}^{q_H} (Iq - \widetilde{q_{home}}) \frac{1}{q_H - q_L} dq\,, \qquad (3.18)$$

Equation (3.17) is differentiable and monotonically increasing in in I, and its first derivative with respect to I equals:

$$\frac{\partial E[s_{market}q_{market}]}{\partial I} = \begin{cases} 0 & if\ I \leq \frac{\widetilde{q_{home}}}{q_H} \\ \frac{1}{2}\frac{(q_H - \widetilde{q_{home}}/I)(q_H + \widetilde{q_{home}}/I)}{q_H - q_L} & if\ \frac{\widetilde{q_{home}}}{q_H} < I \leq \frac{\widetilde{q_{home}}}{q_L}, \\ \mu_q & if\ I > \frac{\widetilde{q_{home}}}{q_L} \end{cases}$$

(3.19)

and is continuous and monotonically increasing in I. Note again that the last condition will never be satisfied if $\widetilde{q_{home}} > q_L$.

For farmers with a high value valuation of safe home consumption β_H, we have $\beta_H\mu_q > c > 0$, so that the optimal investment I^{**} is uniquely defined by:

$$\beta_H \frac{\partial E[s_{home}q_{home}]}{\partial I}\bigg|_{I=I^{**}} + \delta \frac{\partial E[s_{market}q_{market}]}{\partial I}\bigg|_{I=I^{**}} = c, \qquad (3.20)$$

which has a solution for $I \leq 1$, or otherwise $I = 1 > I^*$.

Equation (3.20) can be reduced to:

$$(\beta_H - \delta)\frac{\partial E[s_{home}q_{home}]}{\partial I}\bigg|_{I=I^{**}} + \delta\mu_q = c, \qquad (3.21)$$

The left-hand side of equation (3.21) is decreasing in I^{**} and, by equations (3.15) and (3.20), this exceeds c for $I^{**} \leq I^{*}$. We therefore must have $I^{**} > I^{*}$, meaning that the price premium has a positive effect at the intensive margin of investment. Intuitively, the value of safe harvest in excess of home consumption increases, which causes a marginal increase in the benefits of investments in safe food, leading the farmer to invest more. In this way, the existence of a market premium can increase the investment in food safety and the safety of food consumed by subsistence farmers.

3.3. Study setting

3.3.1. The market for safe maize

The informal markets to which most maize farmers in Kenya sell do not reward unobservable quality (Hoffmann et al. 2013). However, a growing number of maize millers in the formal sector do test for aflatoxin at purchase and reject maize that does not conform to the regulatory standard. These millers offer a significant premium above the spot market price of maize in the informal market.[10] To obtain a premium price, several quality characteristics must typically be met: maize must be at or below 13.5% moisture content; it must conform to grading standards for the proportion of foreign matter, broken, damaged, and discolored kernels; and it must contain total aflatoxins below the regulatory standard. Farmers can meet most of these criteria through adequate drying and removal of sub-standard grains and other particles. The exception is aflatoxin, which may be present without any visible sign of contamination.

The cost of transporting maize from the study region to the Nairobi market, where premium prices can be obtained, is prohibitive in most years. Local millers within the study counties did not screen for aflatoxin at the time of the experiment, though one maize wholesaler did report testing in response to demands by particular buyers, and recently launched a maize flour product.[11] As disposable incomes and concern over food safety grow, and government

[10] On the same day in February 2015, Unga Ltd.'s Eldoret plant was paying 2200 Kenyan Shillings (KSh) – approximately $22 US – for a 90 kg bag of maize, while the price at the open-air market in Eldoret was 1700 KSh.

[11] This product was launched after the conclusion of data collection for the experiment described here, and no buyers were in the market for aflatoxin-safe maize at the time of the experiment.

enforcement of existing regulations strengthens, it is reasonable to expect that that a local premium market will emerge. However, the premium paid by regional millers is likely to be lower than that offered by the miller referenced above, which produces perhaps the two best-known brands in the country, including the most expensive. Nairobi-based millers in the next quality tier offer a premium of between 200-250 KSH per 90 kg bag over the informal market. We propose that a conservative estimate of the premium farmers in the study region could expect to receive from a regional miller for aflatoxin-safe maize, accounting for the lower spending power of consumers in this market, is 100 KSH per bag. This is the aflatoxin safety premium we offer to farmers in the market linkage treatment.

Because the cost of testing for aflatoxin (and other food safety hazards) is high relative to the value of produce sold by the typical smallholder farmer, access to a food safety premium requires that maize is aggregated prior to testing. This can be done through producer groups, which are common in Kenya and throughout sub-Saharan Africa. Such groups are sometimes formed by NGOs or other external actors as a platform for providing agricultural training and extension, or by farmers themselves to aggregate their demand for inputs or their produce and reduce transaction costs or obtain better prices. Farmers in such groups who sell to markets with food safety requirements have a strong incentive to ensure that others in the group treat their fields, analogous to a joint liability lending model.

3.3.2. The technology

Aflasafe™ is a biocontrol product that uses native strains of the Aspergillus fungus that do not produce toxins to outcompete toxigenic strains. Aflasafe™ has been shown in farmer field trials to reduce aflatoxin contamination by between 80% and 99% (Bandyopadhyay et al. 2016). Treatment with Aflasafe™ protects crops throughout the growing cycle and storage period, with no impact on the overall level of fungal colonization or crop yields (Cotty, Antilla, and Wakelyn 2007). Similar aflatoxin biocontrol products have been used on food crops in the United States for over 15 years.

The first African country to register an aflatoxin biocontrol product was Nigeria. There, the main initial adopters of Aflasafe™ have been farmers producing maize used by poultry feed processors (aflatoxin impedes weight gain and increases mortality among poultry). Aflasafe KE01™ was approved by the Kenyan government for general use in June 2015, and domestic manufacturing began in 2017. The cost to produce one kg of Aflasafe™ at scale ranges between US $0.7 and $1.2 depending on currency exchange rates and price of materials

(Bandyopadhyay et al. 2016). Due to the small volume produced in Kenya, the current price of Aflasafe KE01™ is US $1.6.

The mean aflatoxin level in samples collected by one of the authors for a separate study in the same study region in 2015, when aflatoxin contamination was considered moderate, was 17 ppb, 70% higher than the maximum allowable level in Kenya. In 2010, recognized as an aflatoxin outbreak year, the mean level of contamination was 47 ppb, 4.7 times the legal limit (Mutiga et al. 2014). In both years, results from field trials cited above indicate that treating fields with Aflasafe KE™ would have brought the average level of contamination into the legal range.[12]

We set the cost of Aflasafe KE01™ in the study to 80 KSH (US $0.78) per kg; this lies within the range of production costs and takes into account the government of Kenya's expressed support for a partial subsidy targeted to smallholders.[13]

3.3.3. Population and sample

The population for this study consists of maize farmers who are members of existing farmer groups in Meru, Embu and Tharaka Nithi counties, Kenya. The three counties fall in the Eastern region of Kenya, and are known for their high levels of aflatoxin contamination. A list of approximately 250 farmer groups in the study area was acquired through the Cereal Growers' Association (CGA), a national member-based farmer organization, and the Ministries of Agriculture in each of the three counties. From April to August 2017, 224 groups were visited and lists of their members were obtained.[14] From these 224 groups, we selected 152 groups into our experiment.[15]

---

[12] The mean level of contamination (as opposed to the probability of non-compliance for a particular farmer) is relevant both from an economic and health perspective, since most of the health burden of aflatoxin arises though cumulative exposure to moderate levels of the toxin over time, and because maize is tested by processors in large lots.

[13] Together with rainfall insurance, which most farmers purchased when given the choice (and which those offered the Aflasafe™ plus insurance bundle had no choice but to purchase), the cost per kg was 100 KSH (US $0.97).

[14] Some of the groups in the initial list were members of the same Community Based Organization (CBO). In such cases, only one group per CBO was visited for our study.

[15] We selected the 152 groups in a way that minimized the baseline differences in groups assigned to the two insurance conditions. See Hoffmann, Kariuki, et al. (2018) for details.

3.4. Study design

3.4.1. Farmer training and sale of biocontrol product

All 152 groups in our experiment were given information on the benefits of aflatoxin biocontrol and instructions on its use. This was done through two rounds of training, in which all the group members were invited for a half day meeting. The first round of training took place in September-October 2017, planting time in the study area. During these meetings, group members were given information about Aflasafe™ and how rainfall index insurance could be used to insure investment in this technology against weather related shocks. In addition, some of the groups were told they could earn a premium price of 100 KSH per bag of maize grown using Aflasafe™. They were informed that they could only purchase the biocontrol product through the project as it was not available in the study area.

A second round of training was conducted in November and early December, a few weeks after planting and just before the time at which Aflasafe™ should be applied. During these meetings, group members were trained on how to apply Aflasafe™ and how to activate the rainfall insurance offered with the product. A demonstration of Aflasafe™ application was conducted on the farm of one member of each group. At the end of the meeting, those present were given an opportunity to purchase Aflasafe™ and an actuarially fair rainfall index insurance that was specifically designed to insure the investment in Aflasafe™ against weather related shocks. The biocontrol product was offered in packages of 4 kg, a quantity sufficient to treat one acre of land. Farmers who wished to purchase less than 4 kg were requested to pair up with other group members and share a 4 kg package amongst themselves.[16] Farmers who wished to purchase Aflasafe™ later were given a chance to do so through a subsequent sales visit by the project staff.

Both rounds of meetings were conducted by trainers employed by the CGA. CGA trainers had been instructed on the use of Aflasafe™ by the International Institute of Tropical Agriculture (IITA), which supplied the product.

[16] Farmers who paired up were recorded separately, as independent entries in our Aflasafe™ sales data sheets, showing their respective amounts depending on the amount of money paid by each of the farmer.

3.4.2. Experimental design

All 152 groups in the experiment were given the opportunity to purchase Aflasafe™ and trained on its use. Half of these groups were randomly assigned to receive a premium price for safe maize (output market linkage).

During the initial round of training, groups assigned to the output market linkage treatment were promised a bonus of 100 KSH per bag for maize grown using Aflasafe™. The bonus was to be paid shortly after harvest. Members who purchased Aflasafe™ and wanted to sell their maize through the project would aggregate their maize at a central place to be identified by the group members. A rapid qualitative aflatoxin test would be conducted on the aggregated maize to check if the maize had aflatoxin levels higher than the East African limit (10 ppb). Farmers were informed that any aggregated maize that contained levels higher than 10 ppb would not qualify for the bonus. They were advised to record the number of members who purchased Aflasafe™ in their group and the amount purchased by each member, and to ensure that only treated maize was aggregated for testing. Aggregation of maize and payment of the bonus took place in February-March 2018, at the end of the harvest season.

3.4.3. Data

A short survey of all 224 farmer groups on the initial list was conducted during meetings with these groups in April-August 2017 for the purposes of sample selection, stratification, and balance checks. Data on each group's geographical location, as well as their members' familiarity with weather insurance, awareness of aflatoxin, use of agricultural inputs, and levels of maize production and marketing were collected. Lists of the groups' members, indicating whether each member was present during the initial meeting or not, were also obtained.

After selecting 152 groups into the study, baseline survey data was collected from each of these. Baseline data collection took place in September-October 2017, immediately prior to the first training meeting, at the training site. Data was collected both at the individual farmer level and at the farmer group level. This data was used to generate control variables used in the analysis of treatment impact and to further test for balance across treatments. Six farmers per group were randomly selected from among the farmers present during the census meeting. If fewer than six farmers were present at the census meeting, additional farmers were selected from among those listed as members but not present. In case any of the selected farmers were not available, replacements were selected from a randomly ordered list of six additional farmers, selected in the same fashion as the primary sample. A total of 892 individual farmers out of 3605 listed farmers were interviewed.[17] A group level questionnaire was administered to one or more of each group's leaders. A follow-up survey with the same respondents interviewed at baseline was conducted in March-April 2018, at the end of the season. Three of the baseline respondents could not be tracked down for follow-up, resulting in 889 observations in this round of data.

Administrative data on farmers' purchases of Aflasafe™ was collected during sales visits in November and early December 2017. For each farmer who purchased biocontrol (including those who purchased less than 4 kg), name, gender, land area under maize, and the number of packets of Aflasafe™ purchased were recorded. This data was used to construct the main outcome variables: adoption (equal to 1 if the farmer purchased Aflasafe™ and 0 if the farmer did not), and adoption intensity (a continuous variable indicating the amount of the biocontrol product purchased).

3.4.4. Randomization

The 152 groups in our study are located in 124 villages. To eliminate within-village spillover effects, assignment to the market linkage treatment was randomized at the village level. This randomization was stratified by county.[18]

Table 3.1 provides summary statistics by market linkage treatment assignment. For a description of the construction of variables from baseline survey data, we refer to the registered Pre-Analysis Plan (Hoffmann et al. 2017). Rainfall index insurance triggers reflect historic rainfall patterns at the location where group meetings were held.

In general, we find that the market linkage treatment groups are well-balanced on almost all observables. We do, however, find that farmers eligible for the premium price offer were more likely to be present during the census meeting. Given that we test for balance on 27 variables, a significant difference on one of these is not unexpected and does not indicate structural differences across treatments. We control for farmers' presence during the census meeting in the analysis below, as well as for the other observables described in Table 3.1.

---

[17] In 20 groups, it was not possible to interview six farmers and only five were interviewed.

[18] This randomization was also stratified by rainfall index insurance treatment assignment.

Table 3.1. Balance at Baseline across Market Linkage Treatments

	Market linkage			No market linkage			Diff
	N	Mean	SD	N	Mean	SD	P
Bundled insurance	1333	0.552	0.497	1380	0.504	0.500	0.654
Individual present during the census meeting	1333	0.513	0.500	1380	0.423	0.494	0.026
Farmer sex	1333	0.244	0.430	1380	0.214	0.410	0.547
Group mean of:							
- Age of the farmer (completed years)	1333	50.2	7.8	1380	50.1	7.7	0.903
- Years of education completed by head	1333	6.95	2.07	1380	7.03	1.96	0.828
- Relationship with the head	1333	0.601	0.274	1380	0.583	0.294	0.732
- Asset index	1333	5.51	1.07	1380	5.77	1.11	0.224
- Total land under maize main season previous year (acre)	1333	1.49	0.63	1380	1.40	0.77	0.500
- Maize harvest main season previous year (kg)	1333	439	381	1380	424	394	0.838
- Maize marketing: whether sold any maize last season	1333	0.457	0.303	1380	0.489	0.344	0.622
- Total expenditures on agr. inputs & labour main season previous year (KES)	1333	10978	5342	1380	10595	5243	0.682
- Propensity for social learning dummy	1333	0.483	0.257	1380	0.476	0.286	0.898
- Aflatoxin knowledge index	1333	0.012	0.351	1380	-0.063	0.400	0.301
- Knowledge and experience with insurance	1333	1.35	0.42	1380	1.31	0.42	0.601
- Individual trust index	1333	-0.006	0.234	1380	-0.006	0.293	0.989
- Punishment index	1333	0.037	0.471	1380	-0.069	0.522	0.277
- Quantitative risk aversion score	1333	4.43	1.19	1380	4.34	1.29	0.680
County:							
- Meru	1333	0.482	0.500	1380	0.417	0.493	0.542
- Tharaka Nithi	1333	0.167	0.373	1380	0.153	0.360	0.858
- Embu	1333	0.351	0.477	1380	0.430	0.495	0.450
Group level trust index	1333	-0.006	0.234	1380	-0.006	0.293	0.989
Group punishment index	1333	0.037	0.471	1380	-0.069	0.522	0.277
Group capacity index	1333	0.115	0.534	1380	-0.074	0.448	0.121
Proportion of group members female	1333	0.773	0.250	1380	0.800	0.228	0.557
Rainfal index insurance trigger for vegetative stage	1333	35.8	12.0	1380	34.9	13.2	0.740
Rainfal index insurance trigger for flowering stage	1333	1.47	0.60	1380	1.40	0.64	0.619
Rainfal index insurance trigger for ripening stage	1333	94.2	16.1	1380	94.8	16.4	0.862

P-values corrected for village level clustering

3.4.5. Farmer expectations at baseline

Table 3.2 shows summary statistics based on data collected at baseline, of the land farmers planned to plant with maize in the coming season and their expectations of the resulting harvest under normal, poor, and very good conditions. The amount of maize farmers expected to store for household consumption under a normal harvest, and the amount sold (assumed to be any maize not retained for household consumption) are also shown. Note that the amount of maize stored for home consumption in case of a normal harvest is lower than the expected harvest in a poor season, indicating that farmers indeed face considerable upside risk to food safety investments that are not rewarded in the market.

Table 3.2. Baseline Descriptives

	N	Mean	Median	SD
Total land under maize this season (acre)	892	1.68	1.00	1.30
Normal maize harvest this season (kg)	892	925	500	1150
Maize harvest if season is "poor" (kg)	892	367	180	609
Maize harvest if season is "very good" (kg)	891	1431	900	1524
Expected maize harvest this season (kg)	891	1251	900	1338
Amount stored for family consumption from a normal harvest (kg)	892	283	225	213
Amount sold from a normal harvest (kg)	892	630	270	998

Variables winsorized at 99th percentile

Suppose a median farmer would treat half of his one-acre farm with Aflasafe™ with the intention to sell this maize and earn a premium of 100 KSH per 90-kg bag of maize. Dividing the median of farmers' expected harvest under a normal year by the median acreage under maize implies an expected yield of 500 kg per acre, suggesting that the farmer can sell 250 kg of safe maize in a normal season.[19] With a premium of 100 KSH per 90-kg bag, the expected revenue in a normal season is 278 KSH, while the cost of Aflasafe™ (including rainfall insurance) equals 200 KSH[20] plus labor cost. However, in a bad season, the household would need most or all of this maize for home consumption, so that the expected revenue from market sales would be almost zero and the household would make a financial loss.[21] With our modest price premium, the financial return on investment (ROI) for aflatoxin-safe production of maize is thus small and can even be negative.

[19] We estimate mean yield as the ratio of the means rather than the mean of the ratios because some very low values of the area planted combined with large harvests result in implausible values.

[20] We include the cost of insurance, since this was offered to all farmers, and most farmers purchased it.

[21] While the insurance covers part of the investment in bad seasons, the cost of the product is not perfectly insured. While in the season of our study, the harvest was lower than expected for a bad season (see Table 4), farmers who had activated the rainfall index insurance contract received a payout of 59% of the investment in Aflasafe™ on average.

3.5. Empirical strategy for estimation of treatment effects

3.5.1. Main sample

As our experiment allowed new group members to buy Aflasafe™, the final group composition may be endogenous to the experimental treatment. We therefore use the farmers listed during the group census as the sample for analysis.[22]

Since being surveyed at baseline may affect later technology adoption behavior and bias treatment effect estimates in general (Zwane et al. 2011) and in our study specifically (Chapter 2), impacts for the sub-sample of non-surveyed farmers are likely to be most externally valid. We therefore focus our main analysis on farmers that were not surveyed at baseline.[23] In the selection of the survey respondents, preference was given to farmers who were present during the group census meeting. Present farmers are therefore under-represented in this non-surveyed sub-sample. We correct for this under-representation by reweighting observations based on the likelihood of inclusion in the sample, given an individual's presence at the meeting.

3

3.5.2. Treatment effects

To assess the effect of the premium market linkage treatment on farmers' adoption of Aflasafe™, we estimate the following equation both with and without controls:[24]

$$Adoption_{ijv} = \alpha_1 + \beta_1 \cdot Market_v + (\gamma_1 \cdot X_{ij}) + \varepsilon_{ijv1}, \qquad (3.22)$$

where $Adoption_{ijv}$ represents Aflasafe™ adoption or adoption intensity by farmer i in farmer group j in village v, and $Market_j$ indicates whether the group was assigned to the market linkage treatment. X_{ij} is the vector of controls, as specified in the Pre-Analysis Plan (Hoffmann

[22] For one group, the group census list was lost and re-taken later. Although the group size had not changed, this might have affected the sample composition. Excluding this one group from our analysis, however, does not affect our results.

[23] The proportion being surveyed at baseline did not significantly differ across treatment and control groups ($p = 0.627$).

[24] All estimates are intention-to-treat. We cannot estimate the effect of treatment on the treated, as we do not have information on which farmers were aware of the premium price.

et al. 2017) and listed in Table 3.1. ε_{ijv1} is the error term. Standard errors are estimated using the Huber and White sandwich estimator to account for clustering at the village level.

To test the impact of our market linkage treatment, we test whether $\beta_1 = 0$.

3.5.3. Alternative samples

We perform the same analysis on two alternative samples. First, although we randomly selected the sample to be surveyed, actual participation in the survey may be endogenous. Our primary sample, which excludes surveyed farmers, may thus be constituted of relatively less engaged members of participating farmer groups. As a robustness check, we therefore check the representativeness of our results by performing the same analysis as above on the subsample that excludes all twelve farmers who were randomly selected as primary or replacement respondents, while still correcting for the under-representation of present farmers in this sample. Second, we complete the analysis using the full sample, including the five to six farmers surveyed at baseline, as specified in our Pre-Analysis Plan (Hoffmann et al. 2017).[25] As being surveyed at baseline significantly affects adoption in our study, we additionally control for baseline survey status when we include baseline controls in these regressions.

3.6. Results

In Table 3.3 we report estimates of the impact of being offered a premium price. We analyze whether farmers in groups assigned to the market linkage treatment were more likely to purchase any Aflasafe™ and whether adoption intensity, measured as the quantity of Aflasafe™ purchased (in kg), was higher in these groups. Columns (1) and (2) of Table 3.3 show no significant impact of the market linkage treatment on the extensive margin of adoption. The point estimates suggest an increase in adoption of around four percentage points (close to one third of the control group mean), but the estimates are statistically insignificant. In contrast, the amount of Aflasafe™ purchased is significantly higher in groups assigned to the market linkage treatment. The estimates in columns (3) and (4) show that these farmers

[25] The Pre-Analysis Plan also includes equation (3.22) for the sub-sample of surveyed farmers only. The strong impact of being surveyed on Aflasafe™ purchase leads us to conclude that these results do not accurately portray the impact of the interventions. We therefore present the full sample impact results here and leave impact results for the sub-sample of surveyed farmers for another paper (Chapter 2).

purchased 0.28 kg more Aflasafe^{TM} on average, an increase of almost 100% relative to the control group.

Table 3.3. Impact of Market Linkage vs. No Market Linkage

	Outcome variables			
	(1)	(2)	(3)	(4)
	Adoption	Adoption	Intensity (kg)	Intensity (kg)
Market linkage	0.036	0.043	0.276***	0.278***
	(0.026)	(0.027)	(0.083)	(0.078)
Baseline controls	No	Yes	No	Yes
Villages	124	124	124	124
Observations	2713	2713	2713	2713
Mean of no market linkage	0.128	0.128	0.290	0.290

Standard errors clustered at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01

In line with our model, these findings suggest that farmers who purchase Aflasafe^{TM} use it first on maize produced for their own consumption. Indeed, endline descriptive statistics in panel A of Table 3.4 indicate that among farmers who purchased Aflasafe^{TM}, 83% (in the market linkage group) to 89% (in the control group) reported having safe maize for home consumption as a reason for doing so. In contrast, the ability to sell maize at a premium was reported by only 5% of farmers in the control group, compared to 19.4% in the market linkage group. Aflatoxin knowledge increased to the same extent in groups with and without the market linkage (the knowledge index is standardized to have mean zero and standard deviation 1 at baseline), and both groups hold similar beliefs about the efficacy of Aflasafe^{TM}.

Table 3.4. Endline Descriptives

	Market linkage			No market linkage			Diff
	N	mean	sd	N	mean	sd	p
Panel A:							
Aflatoxin knowledge index at endline	448	0.485	0.628	441	0.520	0.579	0.378
Efficacy belief:							
- No chance of aflatoxin	420	0.576	0.495	408	0.600	0.490	0.564
- Small chance of aflatoxin	420	0.276	0.448	408	0.265	0.442	0.741
- 50/50 chance of aflatoxin	420	0.138	0.345	408	0.132	0.339	0.841
Reasons for purchasing Aflasafe (self-reported):							
- To have save maize for home consumption	175	0.829	0.378	176	0.886	0.318	0.156
- To be able to sell my maize at a premium	175	0.194	0.397	176	0.051	0.221	0.000
- To have safe maize for sale	175	0.457	0.500	176	0.438	0.497	0.741
Panel B:							
Among all surveyed farmers:							
- Proportion with positive actual harvest	448	0.902	0.298	441	0.889	0.315	0.657
- Proportion with expected bad harvest < normal year consumption	449	0.559	0.497	443	0.551	0.498	0.887
- Proportion with actual harvest < normal year consumption	448	0.679	0.468	441	0.667	0.472	0.817
Among those with actual harvest < normal year consumption:							
- Actual harvest	304	81.4	108.1	294	67.6	88.0	0.221
- Expected normal harvest	304	804.7	1020.1	294	675.0	803.5	0.261
- Proportion with zero sales	304	0.819	0.386	294	0.891	0.312	0.040
Among those with actual harvest > normal year consumption:							
- Actual harvest	144	573.9	540.7	147	574.5	532.6	0.993
- Expected normal harvest	144	1392.8	1506.4	147	1225.5	1386.9	0.420
- Actual consumption	144	182.5	123.7	147	162.3	110.8	0.200
- Expected normal year consumption	144	247.7	180.6	147	244.4	184.8	0.903
Among those that adopted Aflasafe:							
- Home consumption is safe to the extent possible	132	0.864	0.344	122	0.984	0.128	0.002

P-values corrected for village level clustering

While the premium price did incentivize farmers to purchase more Aflasafe™, it mostly did so among farmers who decided to purchase some Aflasafe™ regardless of the premium.

Our theoretical model predicts that the option of selling at a premium price allowed farmers to invest more in the safety of maize for home consumption in the face of an uncertain harvest. If the harvest was lower than expected, the treated maize would be used for home consumption, while in the case of a bumper crop, the excess could be sold at a premium.

To illustrate, using the mean values shown in Table 3.2, a farmer would need to apply Aflasafe™ to 0.51 acres planted with maize in a normal year to grow a sufficient volume of treated maize for her family's consumption. The cost of Aflasafe™ in this scenario is 206 KSH. But in a bad year, Aflasafe™ would have to be applied to 1.30 acres to attain the same volume of treated maize, at a cost of 519 KSH. Without the market incentive, a farmer might be hesitant to spend this much on Aflasafe™, and risk wasting over 300 KSH in the case of a normal harvest (and even more in case of a good harvest). But with the incentive, such a farmer can safely purchase enough Aflasafe™ to ensure enough treated maize even in a bad year, knowing that if the 1.30 acres yields more maize than her household requires, she will reap a profit of as much as 616 KSH from treating her land, in addition to producing safer maize for her family's consumption.

Our results in Table 3.3 indicate that control group adopters purchased 2.27 kg of Aflasafe™, on average. This is sufficient to treat 0.57 acres of land, very close to the mean value of 0.51 acres required to ensure safe maize for own consumption in a normal year. Adopters in groups receiving the market linkage treatment purchased an average of 3.32 kg, sufficient to treat 0.83 acres. Treating this area of land with Aflasafe™ means a family will be closer to producing enough safe maize to cover their own consumption needs even if the harvest is poor (though still short of the mean poor harvest requirement of 1.3 acres). In effect, the premium makes precautionary investment in the treatment of maize for home consumption less costly.

The season we analyze turned out to be an exceptionally bad one. While Table 3.2 showed that the average (median) farmer expected to harvest 367 (180) kg in case of a poor season, the actual amount of maize harvested was much lower, with a mean of less than 240 kg and a median of 90 kg. This falls short of the amount stored for home consumption in a normal year (see Table 3.2), which allows us to assess some features of our theoretical model.

As reported in Panel B in Table 3.4, the actual harvest fell short of normal year home consumption for two thirds of the farmers in our sample. Among these farmers, the vast majority did not sell any maize.[26]

For farmers whose harvest exceeded normal year home consumption, we see that their actual harvest was about 40% of the expected normal harvest, while actual consumption was about 70% of normal year consumption. While the assumption that farmers consume all harvest up to some amount and sell the remainder is simplifying, we consider it reasonable given these descriptives.

[26] We report the fraction of farmers with zero sales, rather than mean sales, since many farmers reported zero sales and the means are heavily affected by extreme values.

Finally, the bottom row of Table 3.4 shows that most farmers consumed all safe maize at home, and only sold safe maize if the amount of safe maize produced exceeded home consumption. This supports our assumption that farmers first select safe maize for home consumption.

3.6.1. Robustness checks

As discussed in Section 3.5, our main analysis focuses on the sub-sample of non-surveyed farmers. This was driven by the concern that being surveyed at baseline might itself affect technology adoption, and thus bias treatment effects. However, being surveyed at baseline was not entirely random. While we account for the probability of selection given presence at the census meeting through reweighting, we did not always manage to interview the first six sampled farmers. Whether or not a sampled farmer actually participated in the baseline survey could reflect unobserved characteristics correlated with the probability of adopting a new technology.

To assess the robustness of our results with respect to the estimation sample, we estimate treatment effects across two alternative samples. Table A3.1 in the Appendix presents results for the market linkage treatment. In columns (1) to (4) the sample excludes all farmers who were listed for the baseline survey (either in the list of the first six farmers sampled, or in the list of six additional farmers, as explained in Section 3.4.3), irrespective of whether or not they participated in the survey. Columns (5) to (8) present the estimates for the full sample, including all farmers surveyed at baseline. These results indicate that sample selection does not affect our findings: the market linkage premium did not significantly affect Aflasafe™ adoption in either sample, while the amount purchased increased significantly in all samples.

3.7. Discussion

Many food safety hazards, including contamination with fungal toxins, are most effectively addressed at production. Technologies appropriate for use by small-scale producers are available, but adoption is a challenge. In settings where the scale of production is small and output markets informal, incentives to invest in costly to observe attributes such as food safety are absent. To create the market conditions for pass-through of price rewards for food safety, farmers' produce must first be aggregated to a volume at which it can be tested for hazards at reasonable cost.

Another barrier to adoption of food safety technologies is production risk. These technologies must often be applied before the outcome of a stochastic production process is realized. Their use thus increases production costs with certainty, but has an uncertain impact on the value of production. This is the case whether farm produce is consumed solely by the household or sold.

We tested the impact of a price premium for maize that had been treated with the aflatoxin biocontrol product Aflasafe™ and aggregated at the group level on the adoption of this newly available technology. The value of the premium offered for safe maize was modest – approximately 5% of the value of maize in the year it was offered. We find that the group-level premium increases the intensity of Aflasafe™ adoption. This shows that members of producer groups are able to overcome potential barriers to collective action, and trust that others in the group will treat any maize that is aggregated. However, we find no impact on the extensive margin of adoption: farmers that did not adopt Aflasafe™ without the premium are not persuaded to do so when the premium is offered. This finding contrasts with that of a recent study in which a significant increase in adoption was catalyzed by a premium for safe maize at the individual farmer level (Hoffmann and Jones 2018). In that study, the premium was set artificially high, and offered on only the first 45 kg of maize.

Our preferred explanation for this is that farmers who anticipated the possibility of a poor harvest treated a greater area of their farm with Aflasafe™ to ensure themselves of an adequate supply of safe maize for household consumption. This explanation is consistent with the observation that at the end of the maize growing season, which turned out to be a bad one in which rainfall index insurance payouts were triggered for 98% of farmers who had activated a rainfall index insurance contract, only 20 of 76 eligible groups actually aggregated any maize through the project.

Appendix

Table A3.1. Impact of Market Linkage vs. No Market Linkage (Alternative Samples)

	Outcome variables							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Adoption	Adoption	Intensity (kg)	Intensity (kg)	Adoption	Adoption	Intensity (kg)	Intensity (kg)
Market linkage	0.023	0.032	0.231**	0.234***	0.021	0.024	0.215**	0.198**
	(0.027)	(0.027)	(0.090)	(0.082)	(0.027)	(0.026)	(0.089)	(0.078)
Baseline controls	No	Yes	No	Yes	No	Yes	No	Yes
Sample	Non-listed	Non-listed	Non-listed	Non-listed	Full	Full	Full	Full
Villages	118	118	118	118	124	124	124	124
Observations	1795	1795	1795	1795	3605	3605	3605	3605
Mean of no market linkage	0.111	0.111	0.246	0.246	0.174	0.174	0.441	0.441

Standard errors clustered at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01

3

4. The Potency of Private Quality Incentives: Evidence from the Indonesian Dairy Value Chain

Mark Treurniet

Misalignment of quality incentives along value chains may limit smallholder participation in modern value chains. This paper uses survey and administrative data to study how individual quality incentives provided by private actors can help smallholders to improve milk quality. By matching farmers on baseline characteristics, I find that individual quality incentives increased the compositional quality of milk quickly after its introduction. Together with physical inputs and training, individual quality incentives also increased the hygienic quality of milk. Decreasing hygienic quality over time delivered by treated farmers suggests that the impact of the intervention decreased over time.

4.1. Introduction

While food quality standards are increasing, it is not straight-forward that smallholders are included in modern value chains. Increasing quality standards can require investments that are only economically efficient at certain scale, which may be hard to reach for liquidity constrained smallholders (Dolan and Humphrey 2000, Farina et al. 2005, Reardon et al. 2009). Yet, participation in vertical coordinated partnerships has the potential to help smallholders to increase their quality and quantity produced (Swinnen and Maertens 2007). New and cheaper opportunities to measure quality and administer incentives provide the potential to align incentives along value chains by reforming small-scale food transactions, and might help smallholder farmers to further increase quality and keep up with increasing quality standards.

This paper provides evidence that the alignment of incentives along the value chain can lead to substantial smallholder production responses. I study the effects of an intervention in the Indonesian dairy value chain that introduced individual quality incentives for cooperative members. Before the intervention farmers were paid based on the average quality delivered by a group of farmers. In an effort to increase the hygienic quality of milk, part of the cooperative's Milk Collection Points (MCPs) were upgraded, associated farmers were trained and an individual quality incentive system was introduced.

I use a combination of survey data and cooperative administrative data to match farmers that were offered individual quality incentives to similar farmers that were not. A combination of cooperative administrative data and additional quality measures is then used to study the impact on different quality dimensions. The panel structure of the administrative data allows me to study effects over time.

Results show that the intervention led to improved milk quality. As the upgrade of facilities and the training was likely to have significantly affected only the hygienic quality of milk, I estimate the effect on another important quality dimension, compositional quality, to study the isolated effect of individual quality incentives. I find a positive impact on compositional quality, suggesting that individual incentives can be sufficient to improve food quality, even if they are not combined with agricultural training and physical inputs. Together with the physical upgrade and intervention trainings, individual quality incentives also increased hygienic quality. Decreasing hygienic quality at upgraded MCPs suggest that intervention impacts decreased over time as the physical and human capital inputs delivered by the intervention decayed.

This paper contributes to a small, but emerging body of empirical literature on how smallholders respond to market incentives. Generally, incentives are found to matter in the context of agriculture in developing countries (Casaburi and Macchiavello 2015, Bernard et al. 2019, Burchardi et al. 2019). Offering quality incentives specifically requires costly measurement of otherwise often unobservable quality. Depending on who measures quality, and the transparency of the measurement, this introduces a unique form of asymmetric information in which the buyer has more information about quality than the producer (Saenger, Torero, and Qaim 2014, Abate and Bernard 2017).

A number of studies consider impacts of offering quality incentives where quality is measured by a research body or a third party. In most of them, quality is assessed by research bodies and incentives are financed by research funds. Saenger et al. (2013) use a lab-in-the-field experiment to study the effect of offering quality incentives to dairy farmers in Vietnam, and found that both a higher penalty for low quality milk and an extra bonus for high quality increase the use of inputs. Three field experiments found that offering price premiums for conforming to local aflatoxin standards to groundnut farmers in Ghana and maize farmers in Kenya increases the adoption of aflatoxin preventing technologies and decreases aflatoxin contamination (Hoffmann and Jones 2018, Hoffmann, Magnan, et al. 2018, Chapter 3). In one study, quality is assessed by a third party and quality incentives are provided by private actors. Starting from the premise that quality incentives can only be offered if quality is measured, Bernard et al. (2017) randomly provided villages with information on the introduction of scales and quality labelling on local onion markets in Senegal, which allowed sales to be based on weight and quality rather than volume alone. They found that their information treatment increased the use of quality enhancing inputs and quality. Only after the announced scales and labelling were actually introduced with some delay, treated farmers also received significantly higher prices, and earned higher incomes. Yet, all existing papers consider quality measurements by research bodies or external parties, which may be trusted more, but may also involve high transaction costs. To the best of my knowledge, this is the first paper that shows that incentives can also work on the basis of measurements done by the buyer.

On top of that, existing field studies consider effects during at most one agricultural season, which are therefore driven by farmer’s expectations on the production function and quality measurement. By exception, Hoffmann and Jones (2018) study effects during two agricultural seasons to allow some trust to grow after the first season. In contrast, this paper studies impacts over multiple years, during which farmers can monthly update their information after receiving signals from the incentive system.

The remainder of this paper is structured as follows. I first model various mechanisms in which quality can be improved, and then discuss the empirical setting used to test model predictions. After discussing the data, econometrics and results will be discussed by quality dimension. I close with a discussion.

4.2. Model

Since most of the fixed investments needed for quality improvement are provided by the intervention studied, the remaining costs to improve quality are mostly variable and proportional to quantity. I therefore abstract from quantity effects and consider the utility per kg produced. The utility of a farmer i per kg produced is increasing in the price and decreasing in the costs of production:

$$u = \alpha + \beta q^* - \theta c(e_i), \qquad (4.1)$$

where α is the basic price, β is the quality premium, q^* is the observed quality, θ is a scaling parameter for the cost of effort, and $c(e_i)$ is the cost of effort as a function of effort e_i. The cost of effort can be affected by improving physical capital, increasing knowledge and focussing scarce bandwidth (mental capacity and executive control) to quality issues (Schilbach, Schofield, and Mullainathan 2016, Chapter 2).

Observed quality q^* is the weighted mean of the quality of the members of the payment group to which farmer i belongs:

$$q^* = \sum\nolimits_{j \in G} q_j \omega_j\,, \qquad (4.2)$$

where q_j is the quality produced by farmer j, and ω_j is the share of the production of farmer j in the total production of payment group G with $\omega_j \geq 0$ and $\sum_{j \in G} \omega_j = 1$.

Quality q_i is the result of effort:

$$q_i = e_i, \qquad (4.3)$$

Finally, the cost of effort function is both increasing and convex in effort e_i.

Maximizing (4.1) with respect to e_i yields the following first-order condition, which uniquely defines optimal effort e_i^* as long as the farmer produces milk:

$$c'(e_i^*) = \frac{\beta\omega_i}{\theta}, \qquad (4.4)$$

Effort and quality are thus increasing in the ratio $\frac{\beta\omega_i}{\theta}$, which can be interpreted as the size of the effective quality premium relative to the cost of effort parameter. Effort will increase if the quality premium β increases, the share ω_i of farmer's milk in the total group sample increases, or the cost of effort parameter θ decreases.

This results in three hypotheses that will be tested this study:

1. If individual quality incentives replace group incentives, then the share ω_i of farmer's milk in his relevant sample increases to 1, and quality q_i increases.
2. If the quality premium β increases, then quality q_i increases.
3. If the relevant cost of effort parameter θ decreases, then quality q_i increases.

4

4.3. Empirical setting

To test these three theoretical predictions, I study the effects of an intervention at a local dairy cooperative in a peri-urban setting in Indonesia. About 3,700 dairy farmers deliver almost all of their milk twice a day to one of the cooperative's 31 Milk Collection Points (MCPs).[27] The cooperative aggregates the milk and sells most of it to dairy processing companies. In addition, the cooperative provides inputs, extension services, financial services and access to a health facility.

The cooperative faces several challenges in sourcing high quality milk from its members. First, producing good hygienic quality is mainly a behavioral issue. To avoid bacterial contamination, producers should clean and dry the cow's teats before milking, throw away the first milk, which contains most bacteria, filter the milk adequately, and use a proper and clean milk can. While farmers are said to be aware of these principles, the challenge is to have these implemented. Second, compositional quality refers to the mix of chemical constituents, and is

[27] Based on cooperative monitoring data for 2014.

regarded to be higher if the milk contains less water and more solids, like fat, protein and lactose. In general, the compositional quality of milk mainly depends on breed, cow health, feed and added water. While compositional quality varies with cow breeds, broad replacement of cows by another breed is not expected within the scope of this study. Increasing hygienic practices could decrease udder infections and lead to measurable quality improvements over a period of months to a year. Further, most solids are produced when cows are sufficiently fed with an optional combination of easily digestible concentrate and roughage, which takes longer to digest, but promotes a well-functioning digestion. In this specific context, where grasses are scarcely available, the solid content can be gradually increased over a period of weeks to months by increasing roughage intake. Finally, although adulteration with water is said to be rare, the quickest way to increase the solid content of milk is to avoid water to be added accidentally or purposely.

To incentivize members to produce good quality milk, the cooperative measures the quality delivered by groups of 3-8 farmers.[28] Farmers receive a higher price if the quality of their so called payment group is higher. Though using group samples saves testing costs, farmers may not internalize the full benefits of quality improvement, as part of the benefits end up with fellow payment group members.

In an effort to, among other things, increase the hygienic quality of milk and adhere to the Indonesian National Standard (SNI) of 1M bacterial colony-forming units per milliliter (cfu/mL), a public-private partnership (PPP) was initiated between the cooperative, one of its main buyers, another local dairy cooperative and several supporting non-profit organizations, and granted a subsidy of several million euros.[29] Starting in February 2015, the PPP began to upgrade MCP facilities, train associated farmers and implement an individual incentive system

---

[28] Interquartile range in cooperative monitoring data for 2014. The minimum is 1, the median 6, and the maximum 22.

[29] "Development of Sustainable Dairy Villages in Indonesia" is a project of FrieslandCampina Nederland Holding BV, PT Frisian Flag Indonesia, The Friesian Agro Consultancy BV, Stichting Agriterra, Wageningen UR (Stichting Dienst Landbouwkundig Onderzoek), Koperasi Peternak Sapi Bandung Utara (KPSBU) Jabar, Koperasi Peternak Bandung Selantan (KPBS) Pangalengan and the Ministry of Foreign Affairs of the Netherlands through the Facility for Sustainable Entrepreneurship and Food Security (FDOV). Wageningen University is not part of this PPP.

for these farmers. Over a period of three years, the intervention was implemented at seven MCPs.

First, the MCP facilities were upgraded. A new registration and sampling system was installed, and new cleaning facilities for milk cans were built. Farmers also received a loan to buy a new milk can, and a complimentary filter and bucket.

Second, four to eight months before the opening of the new MCP, farmers were invited to attend a socialization meeting in which the upgrading plans were explained. In the following months, farmers attended four to six two-hour trainings on the four hygienic practices discussed above: clean and dry the cow's teats before milking, throw away the first milk, filter the milk adequately, and use a proper and clean milk can. Meanwhile, farmers were visited by extension officers who used checklists to monitor practices, and individual samples were taken to monitor hygienic quality. When more than 95% of the farmers correctly implemented the four hygienic practices, the Total Plate Count (TPC) dropped below 500,000 cfu/mL for all farmers during pre-opening individual testing, and the renovation of the MCP building was completed, the upgraded MCP opened.

Third, starting from the opening of the upgraded MCP, farmers received incentives based on their individual quality. In addition, the TPC was more precisely measured, and extra bonuses were introduced for low levels of TPC.

Since compositional milk quality was not the focus of the intervention, compositional quality was not addressed in trainings and the level of premiums for compositional quality was not changed either. Further, except through long-run effects on cow health, the promoted hygienic practices are unlikely to have substantially affected compositional quality.[30] As a side-effect of the intervention, however, the introduction of an individual incentive system also strengthened incentives to improve compositional quality. I use this dimension of quality to test the first hypothesis and study the isolated effect of individual quality incentives. Effects over time will be studied to shed light on the mechanisms driving quality improvement.

Hygienic quality could be affected in various ways. The extra bonuses increased the quality premium β for hygienic quality that farmers are confronted with. Cost of effort may have

[30] As solids increase during milking, throwing away first milk slightly increases compositional quality, but a large share of the milk should be thrown away to substantially increase compositional quality.

decreased for at least three reasons: (i) farmers received physical capital in the form of buckets, filters and milk can loans, (ii) trainings provided farmers with new knowledge, and (iii) trainings channeled scarce bandwidth (mental capacity and executive control) to the issue of hygienic milk production (Schilbach, Schofield, and Mullainathan 2016; Chapter 2). Finally, introducing an individual incentive system increased ω_i to 1. To jointly test the three hypotheses, I will study whether this bundle of interventions contributes to an increase in the hygienic quality of milk.

Lastly, while behavioral changes can quickly affect hygienic quality, intervention impacts might change over time. On the one hand, the physical and human capital inputs may decay over time. For example, filters might become dirty and buckets might be damaged. And while the hygiene training may temporarily channel bandwidth to the issue of hygienic milk production, the bandwidth channeled to the issue of hygienic milk production may become more thin again after completion of the training. This makes the quality improvement more costly. On the other hand, learning-by-doing might decrease the marginal cost of quality improvement. Whether impacts will decrease or increase over time is thus theoretically ambiguous. I will study which effect dominates by exploring whether impacts on hygienic quality change over time.

4.4. Data

The data used in this paper are derived from three sources: (i) a baseline survey, (ii) cooperative administrative data, and (iii) additional quality tests conducted. A combination of the first two data sources was used to control for pre-existing differences, while a combination of the latter two was used to construct outcome variables.

4.4.1. Baseline variables

To control for socio-economic differences across farmers at upgraded and non-upgraded MCPs, variables were taken from a baseline survey. During October and November 2015 all farmers delivering to 13 selected MCPs, including 6 of the 7 MCPs that would be upgraded, were selected to be visited in their homes for a baseline survey with questions on general demographics, household assets, number of dairy cows, farm assets, labour and trust.[31,32] Out of a sample list of 1351 farmers, 1335 farmers were surveyed (98.8%). Respondents were asked for their member ID, so we could associate baseline survey data with records in cooperative administrative data. When farmers held multiple member IDs, extra member IDs were ignored, as they were more likely to be held temporarily and to be transferred to other farmers during the study period.[33]

To control for the quantity and quality of milk delivered to the cooperative, several indicators were obtained from cooperative administrative data for 2014, the year preceding the start of the intervention, with records for every 10 days and all 3,682 farmers that delivered to the cooperative.[34] The quantity of milk was measured at the individual level for each delivery, and I calculated the sum over the whole year to get the total milk delivered per farmer.[35] To obtain proxies for quality, samples of milk were taken twelve times per month. For budgetary reasons, samples were taken per payment group instead of per individual farmer. These group samples were subsequently tested in the cooperative's laboratory.

For compositional quality, two indicators were available. First, the Total Solids content measures the solid constituents fat, protein and lactose. I used the mean of Total Solids measures over all periods in which the farmer delivered milk as primary indicator for compositional quality. Second, since cows typically produce milk with a freezing point around -0.540°C and adding water raises the freezing point, the Freezing Point of milk delivered can be used as an indicator for added water (Shipe 1959). Farmers received a small penalty on their milk price if the payment group's Freezing Point was above -0.520°C and a larger penalty if it was above -0.500°C. To limit the impact of outliers in the continuous Freezing Point

---

[31] The survey sample was selected at the level of administrative units. Most administrative units had their own MCP building, but some administrative groups shared a MCP building. At one of the MCP buildings, only one of two administrative units was selected.

[32] Details on the construction of survey index variables are included in Table A4.1.

[33] In December 2017 and January 2018, slightly more than two years after the baseline survey, less than 1% of the main IDs was held by a different household, either temporarily or permanently.

[34] Data is missing for June 2014.

[35] To proxy total quantity delivered in 2014, I multiplied the total quantity delivered in the remaining months by the factor 12/11.

variable, I used the frequency of the Freezing Point exceeding these thresholds as secondary proxy for compositional milk quality.

For hygienic quality, one proxy was available. The cooperative performed a Resazurin test, in which the color of milk after a controlled chemical process indicates the degree of bacterial contamination. Farmers received a bonus if the average Resazurin grade of their payment group was supposed to correspond to a TPC below the Indonesian National Standard of 1 million cfu/mL. I used the frequency of this bonus being applied as proxy for hygienic quality.

4.4.2. Outcome variables

Total Solid content and Freezing Points, as indicators for compositional quality, are again taken from cooperative administrative data and are available in panel format until December 2018.[36] While these compositional quality indicators are measured at individual farmer level at upgraded MCPs, they are measured at payment group level at non-upgraded MCPs.

To proxy for hygienic milk quality, I use results from individual sampling and more precise Bactoscan tests, which were introduced at upgraded MCPs and are standard in the global milk processing industry. To obtain similar measures for farmers at non-upgraded MCPs, one individual sample was taken for every farmer that delivered to the non-upgraded MCPs in our baseline survey sample, and also tested with the Bactoscan. These additional quality measures were collected between April and July 2018.

4.4.3. Weighting variable

The quantity of milk produced is taken from cooperative administrative data and available in panel format until December 2018. The quantity is measured at individual level for all farmers.

4.4.4. Sample

The sample consists of cooperative members that (i) delivered at least some milk in 2014, (ii) participated in our baseline survey and (iii) delivered at least some milk that was tested during our endline data collection between February and December 2018. To limit confounding effects of other information acquired, 1 cooperative board member and 16 participants in a small other

---

[36] Until August 2016, record are aggregated per 10 days. As of September 2016, record are aggregated per 15 days. Data is missing for the first half of October 2016.

training program were subsequently excluded from this sample. This leaves a sample of 1064 farmers.

4.4.5. Timing

Table 4.1 shows the timing of data collections together with the implementation of the intervention at the upgraded MCPs in my sample. As can be seen, the baseline survey took place after the opening of the first upgraded MCP and might have been affected by the intervention. Further, the collection of endline quality indicators used in my impact regression analysis started quickly after the opening of the last two upgraded MCPs, begging the issue of longer term impacts. These concerns will be addressed in several robustness checks.

Table 4.1. Timing of Intervention and Data Collection

	2014	2015	2016	2017	2018
MCP upgrades:					
- 1st					
- 2nd					
- 3rd					
- 4th					
- 5th					
- 6th					
Data collection:					
- Baseline administrative					
- Baseline survey					
- Endline hygienic quality					
- Endline compositional quality					

Intervention training periods are indicated in gray

4.5. Impact on compositional quality

I first discuss the empirical strategy and results for the isolated impact of individual quality incentives on compositional milk quality. I start by discussing my matching strategy. I then explain the impact regression models estimated. After studying the evolution of compositional milk quality over time, I conduct some robustness checks.

4.5.1. Weighting

For farmers at non-upgraded MCPs, I only have measures for compositional quality from a group sample, which is a weighted average of individual quality. I will thus compare individual measures from farmers at upgraded MCPs, with weighted group averages from farmers at non-upgraded MCPs. If individual quality is correlated with milk quantity delivered, the ordinary mean deviates from the weighted mean, causing inference in non-weighted regressions to be biased. For the compositional quality, I therefore only show comparisons weighted by milk quantity delivered.

4.5.2. Matching

Appendix Table A4.2 includes baseline summary statistics of milk deliveries to MCPs that were later upgraded and milk deliveries to MCPs that would remain non-upgraded. At the bottom, the table also shows the distribution of available quality measures over months. Farmers at upgraded MCPs were more likely to be female, were more wealthy, had more cows, and had more farm assets. Quality measures were equally available for all months.

To obtain appropriate comparison groups, I employed a Coarsened Exact Matching procedure as discussed by Iacus, King, and Porro (2012). This procedure reweights observations from non-upgraded MCPs to mimic the distribution of observations at upgraded MCPs for pre-selected matching variables. I exactly matched milk on (i) gender,[37] (ii) presence in one of six equidistant intervals of the International Wealth Index, (iii) number of cows, and (iv) number of farm assets.

The exact matching procedure creates very similar matches, but due to the curse of dimensionality causes a substantial part of the sample to remain unmatched. Given that I started with a large sample, a large sample remains, so statistical power does not seriously decrease. Further, since many of these unmatched respondents are not very different from matched ones, the curse of dimensionality does minimally affect external validity. However, the matching

[37] Qualitative research by Wijers (2019) indicates that while women are generally important actors in smallholder milk production in Indonesia, their roles are often not formalised within the cooperative. As women who are formally responsible for dairy farming might come from households with a different intra-household work division, I control for gender in my analysis.

procedure may drop respondents, whose milk still counts in the group measure of others, and this issue will be addressed in a robustness check.

Since quality measures are equally available over time, there is no need to control for differential seasonal effects. Yet, for consistency with later analysis, I corrected for potential seasonal effects by calculating for each farmer a weighted average over available quality indicators, with weights for farmers at to-be-upgraded MCPs chosen such that the distribution over months for farmers at intervention MCPs equals the distribution for farmers at comparison MCPs.

Table 4.2 shows balance at baseline after matching. The intervention groups are now very similar. There is no difference on variables directly used in the matching procedure, while differences for other variables are limited.

4.5.3. Impact regressions

Table 4.2 indicates that non-compliance by farmers is limited. Within the matched sample, both switching from intervention MCPs to comparison MCPs and vice versa was rare. To correct for these limited individual selection effects, I estimated Intention-To-Treat (ITT) effects by using as intervention indicator the actual endline upgrade status of the MCP to which the farmer delivered its last milk in 2014.

To study the effect of individual incentives on compositional quality, I regressed the Total Solids (TS) content and the frequency of the Freezing Point (FP) being above -0.520°C on the intervention indicator.[38] As control variables, I included the baseline characteristics that were listed in Table 4.2, including baseline indicators for both compositional quality variables. Results can be interpreted as average causal impact of the intervention at upgraded MCPs (Average Treatment Effect on the Treated, ATT) under the assumption that matched famers at non-upgraded MCPs provide an accurate counterfactual for farmers at upgraded MCPs.

[38] I do not show results for the frequency of the Freezing Point being above -0.500°C, as this occurred very rarely at both upgraded and non-upgraded MCPs.

Table 4.2. Balance at Baseline after Matching (Comparing Milk across Intervention Status)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	381	44.10	12.01	374	44.15	10.18	0.978
Gender of respondent	381	0.072	0.259	374	0.072	0.259	1.000
Junior high school or higher	381	0.297	0.458	374	0.304	0.461	0.960
Household asset index	381	4.072	1.540	374	4.401	1.653	0.168
International Wealth Index	381	67.62	10.60	374	67.24	10.77	0.842
Progress out of Poverty Index	381	38.93	7.92	374	39.18	8.66	0.878
Number of dairy cattle total	381	4.862	2.243	374	4.862	2.243	1.000
Farm asset index	381	2.738	0.617	374	2.738	0.617	0.957
Number of non-family fulltime workers	381	0.049	0.267	374	0.024	0.160	0.236
Number of non-family parttime workers	381	0.029	0.217	374	0.004	0.067	0.146
Number of family fulltime workers	381	1.130	0.685	374	1.226	0.613	0.150
Number of family parttime workers	381	0.642	0.531	374	0.613	0.505	0.682
Milk quantity (1,000 L)	381	10.235	6.194	374	9.896	5.468	0.788
Milk quantity including extra IDs (1,000 L)	381	10.521	6.693	374	10.240	6.039	0.794
Milk compositional quality TS (%)	381	11.76	0.28	374	11.76	0.27	0.972
Milk compositional quality FP (>-0.520°C)	381	0.172	0.209	374	0.189	0.227	0.694
Milk compositional quality FP (>-0.500°C)	381	0.012	0.040	374	0.007	0.036	0.520
Milk hygienic quality TPC (bonus)	381	0.553	0.280	374	0.610	0.306	0.454
Milk price (1,000 IDR)	381	4.153	0.147	374	4.155	0.145	1.000
Milk income (1,000,000 IDR)	381	39.08	23.89	374	37.79	21.04	0.790
Milk income including extra IDs (1,000,000 IDR)	381	40.17	25.77	374	39.12	23.36	0.800
Payment group size	381	6.245	4.111	374	6.602	3.492	0.824
Distance to MCP (m)	381	857.8	820.4	374	784.2	829.3	0.812
Trust index[1]	370	3.619	0.367	371	3.667	0.345	0.314
Children want to take over	381	0.703	0.344	374	0.667	0.348	0.532
Upgraded at endline	381	0.993	0.086	374	0.010	0.100	0.000
Weight quality measures Feb-2018	381	0.091	0.019	374	0.091	0.023	0.992
Weight quality measures Mar-2018	381	0.091	0.016	374	0.091	0.019	0.976
Weight quality measures Apr-2018	381	0.091	0.013	374	0.091	0.020	1.000
Weight quality measures May-2018	381	0.091	0.009	374	0.091	0.017	0.980
Weight quality measures Jun-2018	381	0.091	0.009	374	0.091	0.016	0.980
Weight quality measures Jul-2018	381	0.091	0.011	374	0.091	0.017	0.973
Weight quality measures Aug-2018	381	0.091	0.011	374	0.091	0.016	1.000
Weight quality measures Sep-2018	381	0.091	0.010	374	0.091	0.016	1.000
Weight quality measures Oct-2018	381	0.090	0.012	374	0.090	0.046	1.000
Weight quality measures Nov-2018	381	0.091	0.012	374	0.091	0.017	0.998
Weight quality measures Dec-2018	381	0.090	0.014	374	0.090	0.047	0.990

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

4

Table 4.3 shows impact regression results. In line with hypothesis 1, both indicators reveal a significant improvement. Estimates suggest that the Total Solids content increased on average by 3.5-3.8% relative to the counterfactual mean, which corresponds to 1.39-1.54 standard deviations. Since processors generally have no value for water content, this suggests that the value of milk increased by 3.5-3.8%. The frequency of the Freezing Point being above -0.520°C almost decreased to zero, which confirms that the water content was lower.

Table 4.3. Impact on Compositional Quality

	Outcome variables			
	(1)	(2)	(3)	(4)
	TS (%)	TS (%)	FP>-0.520°C	FP>-0.520°C
Upgraded (ITT)	0.410***	0.453***	-0.0344***	-0.0348***
- Clustered SE	(0.080)	(0.070)	(0.0103)	(0.0101)
- Bootstrapped p-value	0.000	0.000	0.010	0.004
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	13	13	13	13
Observations	755	755	755	755
Mean of non-upgraded	11.870	11.870	0.0346	0.0346
SD of non-upgraded	0.295	0.295	0.0837	0.0837

Table shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

4.5.4. Mechanisms

Studying effects of the individual quality incentives over time can shed light on the likely mechanisms that farmers have used to increase compositional quality.

Figures 4.1 and 4.2 show the evolution of compositional quality indicators over time for farmers at upgraded MCPs and their comparison groups, broken down by upgraded MCP.[39] Although trends in Total Solids content differ across upgraded MCPs, a substantial and quick

[39] Table 4.2 indicated that a very small proportion of farmers switched from to-be-upgraded MCPs to not-to-be-upgraded MCPs. While these farmers are included in the Intention-to-treat impact regressions, they are not shown in Figures 4.1 and 4.2.

4

increase in the Total Solids content is observed at all upgraded MCPs around their opening. Over the years, the frequency that the Freezing Point exceeded -0.520°C decreased, but only after the upgrade of MCPs, Freezing points above -0.520°C virtually disappeared for a longer period of time.

Given the quick responses in compositional quality, the increase in compositional quality is most likely explained by decreasing added water by about 3.5-3.8%. As impacts are not observed to increase over time, improved feed and cow health are not likely to have substantially contributed to better compositional quality. The impacts found on compositional quality are thus unlikely to be a by-product of the intervention's efforts to increase the hygienic quality of milk, but instead the result of introducing individual incentives for compositional quality.

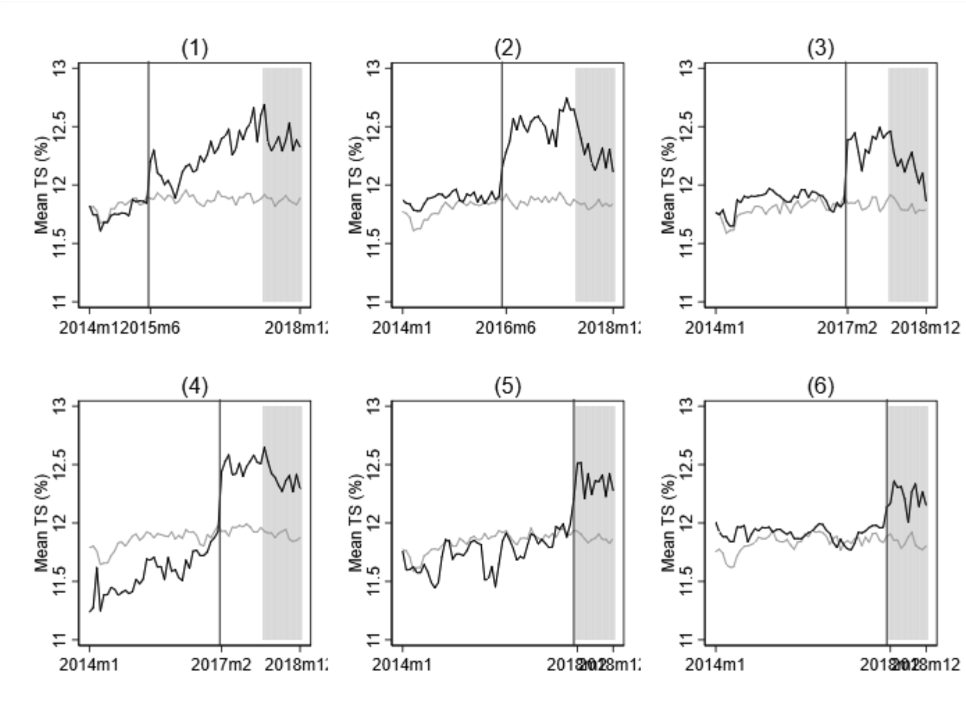

Figure 4.1. Evolution of Total Solids (TS) content of farmers at upgraded MCPs and matched farmers at non-upgraded MCPs over time. Grey shaded months indicate measures used in the impact regressions. Vertical lines indicate the opening of the upgraded MCP

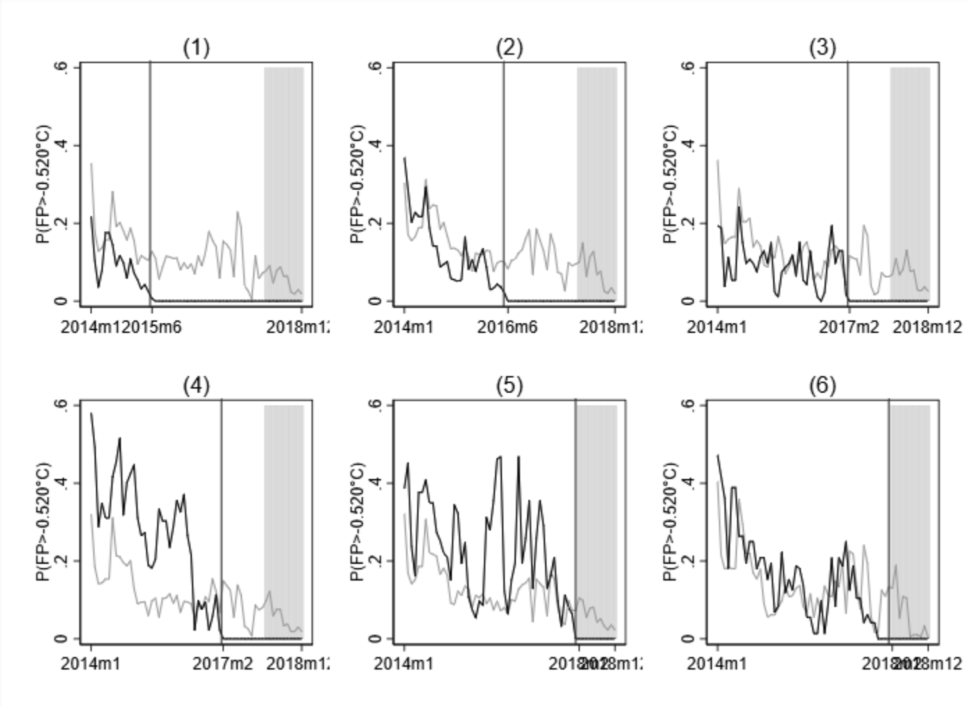

Figure 4.2. Evolution of the frequency of the Freezing Point (FP) exceeding -0.520°C of farmers at upgraded MCPs and matched farmers at non-upgraded MCPs over time. Grey shaded months indicate measures used in the impact regressions. Vertical lines indicate the opening of the upgraded MCP

4.5.5. Robustness checks

Although the selection of matching variables directly follows from the balance table, one might wonder whether the regression results are caused by my matching procedure. As a robustness check, I therefore re-estimated the main results in Table 4.3 while skipping the matching stage. Results are included in Appendix Table A4.3 and are similar to the main results.

As the baseline survey was held after the opening of the first upgraded MCP, one might fear that some baseline survey characteristics are affected by the intervention, thus invalidating the construction of comparable intervention groups. As a robustness check, I therefore repeated my empirical procedure while excluding the first upgraded MCP.[40,41] Results are included in Appendix Table A4.4 and are again similar to the main results.

To study whether impacts remain after a somewhat longer period, I repeated my empirical procedure while excluding the last two upgraded MCPs, which opened in 2018. Results are included in Appendix Table A4.5 and are very similar to the main results.

Finally, one might fear that missing group members and matching invalidate inference on group-level quality measures. To address this concern, I excluded incomplete groups, which had one or more members missing from the matched sample, either because (i) they did not deliver milk in the baseline year 2014, (ii) they did not participate in the baseline survey or (iii) no good match was found. And instead of matching farmers at non-upgraded MCPs to farmers at upgraded MCPs, I matched farmers at upgraded MCPs to farmers at non-upgraded MCPs, so that the sampling weights for farmers from non-upgraded MCPs are not interacted with matching weights. Results are included in Appendix Table A4.6 and are again similar.

4.6. Impact on hygienic milk quality

I now turn to the empirical strategy and results for the impact on hygienic quality. As discussed, I study the combined effect of the physical upgrade, intervention trainings and individual quality incentives, and how this effect changes over time. The structure mimics the structure of the previous section, and the discussion focuses on deviations from the strategy employed before.

4.6.1. Weighting

As I have individual measures for hygienic quality at endline, I am able to show both farmer-level comparisons as well as milk-level comparisons. While in the main text a balance table is only shown for the farmer-level analysis, I also present impact regressions results of the

---

[40] For every re-matching, balance tables before and after matching are available on request.

[41] For this and some other robustness checks holds that after matching, the number of non-family full-time workers was found to be significantly larger for farmers from upgraded MCPs, while the absolute difference was still small. If anything, principal-agency challenges might make it harder to increase quality for farmers with non-family workers.

analysis weighted by milk quantity, for comparison with the previous section and because of relevance for policy. Farmer-level results do not structurally differ from milk-level results.

4.6.2. Matching

Appendix Table A4.7 includes baseline summary statistics of milk deliveries to MCPs that were later upgraded and milk deliveries at MCPs that would remain non-upgraded. Although the samples slightly differ due to data availability, the sample is quite similar as before, causing the same patterns to be observed. Farmers at non-upgraded MCPs were slightly more likely to have their quality being tested in later months, although differences are not statistically significant.

To obtain appropriate comparison groups, I employed the same Coarsened Exact Matching procedure with the same matching variables as before. I also used the same strategy to correct for potential seasonal effects.

Table 4.4 shows balance at baseline after matching. The intervention groups are again very similar.

Table 4.4. Balance at Baseline after Matching (Comparing Farmers across Intervention Status)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	356	44.81	12.49	343	44.79	10.60	0.930
Gender of respondent	356	0.070	0.256	343	0.070	0.256	0.984
Junior high school or higher	356	0.267	0.443	343	0.301	0.460	0.724
Household asset index	356	3.944	1.485	343	4.274	1.598	0.144
International Wealth Index	356	66.42	10.86	343	66.04	10.80	0.760
Progress out of Poverty Index	356	38.51	8.28	343	38.51	8.80	0.994
Number of dairy cattle total	356	4.430	2.138	343	4.430	2.138	1.000
Farm asset index	356	2.697	0.622	343	2.697	0.622	0.974
Number of non-family fulltime workers	356	0.053	0.271	343	0.022	0.171	0.106
Number of non-family parttime workers	356	0.011	0.130	343	0.004	0.060	0.688
Number of family fulltime workers	356	1.135	0.662	343	1.204	0.585	0.308
Number of family parttime workers	356	0.626	0.534	343	0.592	0.511	0.676
Milk quantity (1,000 L)	356	8.715	5.558	343	9.014	5.188	0.750
Milk quantity including extra IDs (1,000 L)	356	8.894	5.848	343	9.288	5.722	0.676
Milk compositional quality TS (%)	356	11.72	0.29	343	11.74	0.27	0.868
Milk compositional quality FP (>-0.520°C)	356	0.212	0.237	343	0.215	0.241	0.964
Milk compositional quality FP (>-0.500°C)	356	0.018	0.052	343	0.008	0.033	0.326
Milk hygienic quality TPC (bonus)	356	0.511	0.287	343	0.576	0.308	0.416
Milk price (1,000 IDR)	356	4.127	0.156	343	4.141	0.147	0.754
Milk income (1,000,000 IDR)	356	33.09	21.27	343	34.32	19.96	0.740
Milk income including extra IDs (1,000,000 IDR)	356	33.78	22.39	343	35.38	22.12	0.648
Payment group size	356	6.204	3.970	343	6.800	3.500	0.752
Distance to MCP (m)	356	902.2	834.8	343	788.7	809.9	0.674
Trust index[1]	346	3.605	0.375	340	3.648	0.356	0.324
Children want to take over	356	0.691	0.356	343	0.656	0.360	0.456
Upgraded at endline	356	1.000	0.000	343	0.014	0.116	0.000
Weight quality measures Apr-2018	356	0.260	0.253	343	0.260	0.437	1.000
Weight quality measures May-2018	356	0.300	0.295	343	0.300	0.457	1.000
Weight quality measures Jul-2018	356	0.439	0.305	343	0.439	0.494	1.000

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

4.6.3. Impact regressions

To study the effect of the intervention on hygienic quality, I regressed the TPC and the frequency of the TPC being below the Indonesian National Standard of 1M cfu/mL on the intervention indicator. I included the same control variables as before, which include the baseline indicator for TPC. Since knowledge was likely to spillover to non-upgraded MCPs via regular activities of cooperative extension officers, my results underestimate effects of the knowledge component of the intervention, and are therefore a conservative estimate of the total intervention impact.

Table 4.5 shows impact regression results. The hygienic quality significantly improved, as the TPC decreased by about one third, and the average TPC at upgraded MCPs now satisfies the national standard. The TPC of individual deliveries was also significantly more likely to be below the national standard. Milk-level results are similar to the results of the farmer-level comparison.

Table 4.5. Impact on Hygienic Quality

	Outcome variables			
	(1)	(2)	(3)	(4)
	TPC	TPC	TPC<1M	TPC<1M
Upgraded (ITT)	-414318.9**	-404451.1**	0.134***	0.136***
- Clustered SE	(133605.0)	(106018.1)	(0.037)	(0.028)
- Bootstrapped p-value	0.016	0.014	0.000	0.002
Baseline controls	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer
MCPs	13	13	13	13
Observations	699	699	699	699
Mean of non-upgraded	1177466.5	1177466.5	0.709	0.709
SD of non-upgraded	1516705.6	1516705.6	0.455	0.455
Upgraded (ITT)	-420193.9**	-384799.9***	0.126***	0.121***
- Clustered SE	(132503.6)	(66827.2)	(0.039)	(0.027)
- Bootstrapped p-value	0.014	0.002	0.006	0.000
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	13	13	13	13
Observations	699	699	699	699
Mean of non-upgraded	1126619.7	1126619.7	0.727	0.727
SD of non-upgraded	1461142.4	1461142.4	0.446	0.446

Top panel shows comparison at farmer level, bottom panel shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

4.6.4. Effects over time

Since we have only one TPC measure per farmer at non-upgraded MCPs, the study of evolution of hygienic quality over time is limited to upgraded MCPs. Figure 4.3 shows the evolution of hygienic quality for all farmers in the matched sample per upgraded MCP.[42] The hygienic quality decreased and the TPC increased over time after the opening of the upgraded MCP, suggesting that the impact of the intervention on hygienic quality decreased over time. As this was seen as a problem by the processor and the cooperative, extension officers again visited all farmers with checklists. They frequently found that farmers used dirty filters or inappropriate buckets. Addressing these problems would be easy and cheap, but had not received sufficient attention of farmers. As the period of endline quality indicators used in my impact regression analysis starts soon after the opening of the last two upgraded MCPs, one might wonder whether impacts remain significantly when those two MCPs would be excluded from the analysis. This issue will be addressed in a robustness check.

[42] Breakdown by MCP based on the upgraded MCP to which the farmer delivered first. I corrected for sample selection effects caused by missing values by replacing missing values by the most recent available measure (or the first available measure if no earlier measure was available).

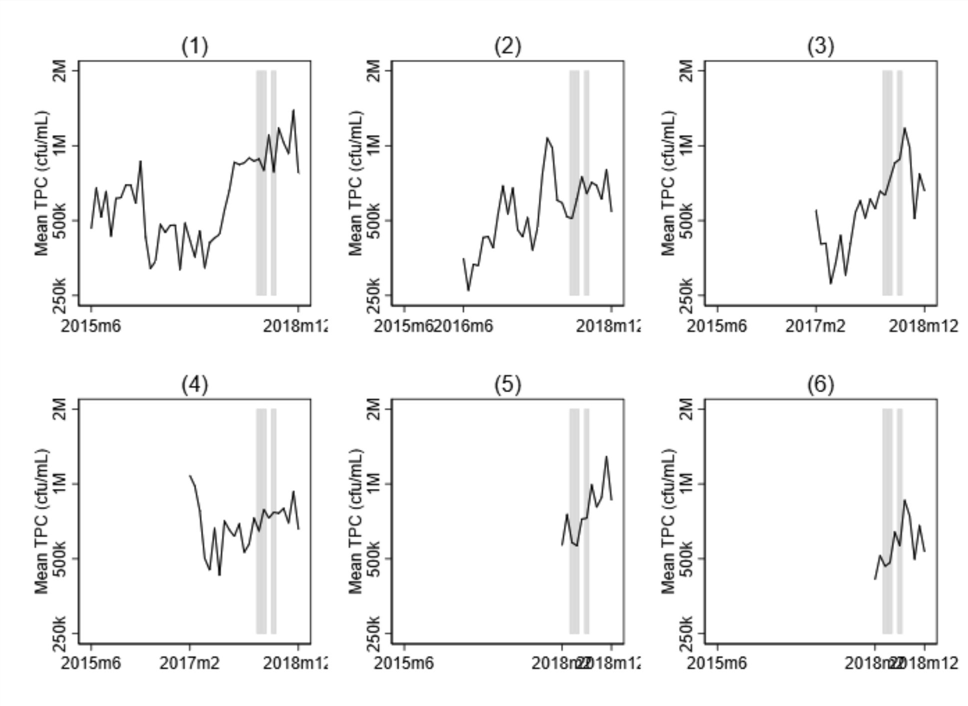

Figure 4.3. Evolution of TPC of farmers at upgraded MCPs over time. Grey shaded months indicate measures used in the impact regressions.

4.6.5. Robustness checks

Appendix Table A4.8 shows that similar results are found when skipping the matching stage, suggesting that impact results do not critically depend on the matching stage.

If farmers from the first upgraded MCP are excluded, results are still significant[43] and point estimates are somewhat larger (see Appendix Table A4.9). If farmers from the MCPs that were upgraded in 2018 are excluded, results are still significant and point estimates are somewhat smaller (see Appendix Table A4.10). Both results are in line with the suggestion that the impact of the intervention on hygienic quality decreased over time.

[43] In the milk-level impact regression of the binary TPC measure with controls included, the intervention indicators is significant only at the 10% level. The larger point estimate suggests that this is mainly caused by decreased power.

4.7. Discussion

While quality standards are increasing, smallholder farmers may or may not be included in modern value chains. New and cheaper opportunities to measure quality have the potential to reform small-scale transactions and might help farmers to keep up with increased quality standards.

This paper finds that the introduction of an individual quality incentive system at a local dairy cooperative in Indonesia increased the compositional quality of milk. Together with physical inputs and training, individual quality incentives also increased the hygienic quality of milk. Results thus confirm that price incentives are a potent tool for quality improvement.

From the moment that the intervention studied in this paper started as a subsidized pilot, the individual incentive system has been maintained by the dairy processor and local cooperative already for years. Two factors may have contributed to this sustainability. First, once investments to implement the system were made, maintaining the system seemed to be in the economic interest of all value chain actors involved, and the value chain did not involve traders that would benefit from non-transparent transactions, as was the case in the study of Bernard et al. (2017). Second, while in the past the cooperative had cancelled penalties for high levels of TPC as members protested against the low prices that resulted for low quality milk, the intervention meetings stressed that farmers could use the intervention to earn higher prices. Moreover, as quality increased at upgraded MCPs, and the price function was not differentiated between upgraded and non-upgraded MCPs, farmers at upgraded MCPs earned higher prices than farmers at non-upgraded MCPs. This may have led cooperative members to accept potential low prices that can result from the individual incentive system.

Appendix

Table A4.1. Details on Survey Variables

Household asset index	Sum of dummies indicating whether the houshold owns at least one:
	- Watch
	- Mobile telephone
	- Smart-phone
	- Bank account
	- Radio
	- Television
	- Refridgerator
	- Freezer
	- Computer
	- Bicycle
	- Motor
	- Car of truck
	- Generator
	- Solar panel
	- Gas cilinder
International Wealth Index	See Smits and Steendijk (2015)
Progress out of Poverty Index	See Schreiner (2012)
Farm asset index	Sum of dummies indicating whether the houshold owns at least one:
	- Barn or cowshed
	- Chopper for cutting the grass
	- Animal-drawn cart
	- Milk can
	- Milking machine
	- Irrigation equipment
Trust index	Mean of 5-point Likert scores on trust in:
	- The local dairy cooperative
	- The local government
	- The processor that was part of the PPP
	- The other processor that buys a lot of milk from the cooperative
	- Other dairy farmers in your payment group
	- Other dairy farmers in your farm group
	- Other farmers in general
Children want to take over	Do you think your children would want to join or take over your dairy farming business at some stage?
	- 0 if No
	- 0.5 if Uncertain
	- 1 if Yes

Table A4.2. Balance at Baseline before Matching (Comparing Milk across Intervention Status)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	583	43.12	11.20	481	45.39	11.56	0.036
Gender of respondent	583	0.174	0.379	481	0.084	0.278	0.002
Junior high school or higher	583	0.361	0.481	481	0.293	0.456	0.518
Household asset index	583	4.708	1.937	481	4.214	1.695	0.072
International Wealth Index	583	71.80	13.03	481	63.33	13.31	0.004
Progress out of Poverty Index	583	40.77	8.83	481	39.12	9.51	0.520
Number of dairy cattle total	583	7.630	5.341	481	5.269	3.943	0.016
Farm asset index	583	2.948	0.775	481	2.204	0.771	0.008
Number of non-family fulltime workers	583	0.173	0.544	481	0.143	0.665	0.724
Number of non-family parttime workers	583	0.051	0.315	481	0.008	0.091	0.110
Number of family fulltime workers	583	1.262	0.841	481	1.251	0.712	0.894
Number of family parttime workers	583	0.625	0.550	481	0.629	0.557	0.896
Milk quantity (1,000 L)	583	16.089	13.842	481	11.880	10.422	0.180
Milk quantity including extra IDs (1,000 L)	583	16.351	13.980	481	12.326	11.075	0.202
Milk compositional quality TS (%)	583	11.76	0.27	481	11.70	0.27	0.498
Milk compositional quality FP (>-0.520°C)	583	0.164	0.202	481	0.200	0.224	0.330
Milk compositional quality FP (>-0.500°C)	583	0.011	0.037	481	0.008	0.036	0.570
Milk hygienic quality TPC (bonus)	583	0.569	0.276	481	0.578	0.303	0.896
Milk price (1,000 IDR)	583	4.161	0.140	481	4.136	0.146	0.428
Milk income (1,000,000 IDR)	583	61.48	52.95	481	45.31	40.45	0.168
Milk income including extra IDs (1,000,000 IDR)	583	62.48	53.48	481	47.03	43.02	0.204
Payment group size	583	5.690	3.983	481	7.094	4.156	0.446
Distance to MCP (m)	583	762.6	775.1	481	642.5	857.5	0.672
Trust index[1]	567	3.615	0.358	478	3.676	0.341	0.122
Children want to take over	583	0.733	0.352	481	0.696	0.325	0.342
Upgraded at endline	583	0.996	0.064	481	0.022	0.142	0.000
Weight quality measures Feb-2018	583	0.091	0.016	481	0.091	0.017	0.892
Weight quality measures Mar-2018	583	0.091	0.014	481	0.091	0.013	0.154
Weight quality measures Apr-2018	583	0.091	0.012	481	0.091	0.015	0.748
Weight quality measures May-2018	583	0.091	0.008	481	0.092	0.012	0.120
Weight quality measures Jun-2018	583	0.091	0.008	481	0.092	0.010	0.184
Weight quality measures Jul-2018	583	0.091	0.009	481	0.091	0.011	0.516
Weight quality measures Aug-2018	583	0.091	0.009	481	0.091	0.011	0.812
Weight quality measures Sep-2018	583	0.091	0.009	481	0.091	0.010	0.434
Weight quality measures Oct-2018	583	0.091	0.011	481	0.090	0.012	0.746
Weight quality measures Nov-2018	583	0.091	0.010	481	0.090	0.013	0.016
Weight quality measures Dec-2018	583	0.090	0.012	481	0.090	0.014	0.298

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

4

Table A4.3. Simple OLS Regressions Compositional Quality

	Outcome variables			
	(1)	(2)	(3)	(4)
	TS (%)	TS (%)	FP>-0.520°C	FP>-0.520°C
Upgraded (ITT)	0.363***	0.372***	-0.0407***	-0.0400***
- Clustered SE	(0.056)	(0.061)	(0.0074)	(0.0087)
- Bootstrapped p-value	0.000	0.000	0.002	0.002
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	13	13	13	13
Observations	1064	1064	1064	1064
Mean of non-upgraded	11.889	11.889	0.0408	0.0408
SD of non-upgraded	0.253	0.253	0.0961	0.0961

Table shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

Table A4.4. Impact on Compositional Quality (1st Upgraded MCP Excluded)

	Outcome variables			
	(1)	(2)	(3)	(4)
	TS (%)	TS (%)	FP>-0.520°C	FP>-0.520°C
Upgraded (ITT)	0.360***	0.404***	-0.0364**	-0.0377**
- Clustered SE	(0.075)	(0.076)	(0.0105)	(0.0113)
- Bootstrapped p-value	0.004	0.002	0.012	0.012
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	12	12	12	12
Observations	653	653	653	653
Mean of non-upgraded	11.881	11.881	0.0367	0.0367
SD of non-upgraded	0.285	0.285	0.0859	0.0859

Table shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

4

Table A4.5. Impact on Compositional Quality (MCPs Upgraded in 2018 Excluded)

	Outcome variables			
	(1)	(2)	(3)	(4)
	TS (%)	TS (%)	FP>-0.520°C	FP>-0.520°C
Upgraded (ITT)	0.403***	0.455***	-0.0343**	-0.0355***
- Clustered SE	(0.086)	(0.076)	(0.0104)	(0.0108)
- Bootstrapped p-value	0.002	0.000	0.014	0.006
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	11	11	11	11
Observations	659	659	659	659
Mean of non-upgraded	11.876	11.876	0.0346	0.0346
SD of non-upgraded	0.300	0.300	0.0853	0.0853

Table shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

Table A4.6. Impact on Compositional Quality (ATU, Incomplete Groups Excluded)

	Outcome variables			
	(1)	(2)	(3)	(4)
	TS (%)	TS (%)	FP>-0.520°C	FP>-0.520°C
Upgraded (ITT)	0.412***	0.473***	-0.0517***	-0.0523***
- Clustered SE	(0.080)	(0.069)	(0.0142)	(0.0156)
- Bootstrapped p-value	0.002	0.000	0.006	0.004
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	13	13	13	13
Observations	457	457	457	457
Mean of non-upgraded	11.887	11.887	0.0517	0.0517
SD of non-upgraded	0.286	0.286	0.1091	0.1091

Table shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

Table A4.7. Balance at Baseline before Matching (Comparing Farmers across Intervention Status)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	562	44.05	12.23	439	45.13	11.85	0.242
Gender of respondent	562	0.169	0.375	439	0.091	0.288	0.002
Junior high school or higher	562	0.327	0.470	439	0.276	0.447	0.506
Household asset index	562	4.301	1.785	439	3.973	1.582	0.082
International Wealth Index	562	68.46	12.87	439	61.78	12.77	0.028
Progress out of Poverty Index	562	39.69	8.81	439	37.47	9.04	0.104
Number of dairy cattle total	562	5.893	4.300	439	4.223	2.835	0.010
Farm asset index	562	2.835	0.789	439	2.200	0.817	0.076
Number of non-family fulltime workers	562	0.100	0.406	439	0.052	0.344	0.200
Number of non-family parttime workers	562	0.025	0.214	439	0.005	0.067	0.218
Number of family fulltime workers	562	1.221	0.763	439	1.205	0.661	0.848
Number of family parttime workers	562	0.625	0.547	439	0.595	0.549	0.642
Milk quantity (1,000 L)	562	10.900	9.162	439	8.838	6.575	0.144
Milk quantity including extra IDs (1,000 L)	562	11.075	9.330	439	9.158	7.347	0.192
Milk compositional quality TS (%)	562	11.73	0.29	439	11.69	0.27	0.648
Milk compositional quality FP (>-0.520°C)	562	0.197	0.230	439	0.228	0.238	0.542
Milk compositional quality FP (>-0.500°C)	562	0.016	0.048	439	0.011	0.044	0.528
Milk hygienic quality TPC (bonus)	562	0.529	0.289	439	0.538	0.313	0.910
Milk price (1,000 IDR)	562	4.138	0.151	439	4.122	0.153	0.682
Milk income (1,000,000 IDR)	562	41.48	35.02	439	33.55	25.41	0.146
Milk income including extra IDs (1,000,000 IDR)	562	42.16	35.68	439	34.78	28.45	0.208
Payment group size	562	5.959	3.888	439	7.227	3.937	0.516
Distance to MCP (m)	562	840.0	811.6	439	723.8	917.9	0.636
Trust index[1]	546	3.604	0.366	436	3.672	0.355	0.106
Children want to take over	562	0.713	0.358	439	0.680	0.337	0.376
Upgraded at endline	562	1.000	0.000	439	0.023	0.149	0.000
Weight quality measures Apr-2018	562	0.270	0.237	439	0.192	0.391	0.672
Weight quality measures May-2018	562	0.208	0.219	439	0.219	0.412	0.968
Weight quality measures Jul-2018	562	0.521	0.265	439	0.589	0.489	0.758

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

4

Table A4.8. Simple OLS Regressions Hygienic Quality

	Outcome variables			
	(1)	(2)	(3)	(4)
	TPC	TPC	TPC<1M	TPC<1M
Upgraded (ITT)	-443412.8**	-402924.8***	0.166***	0.156***
- Clustered SE	(94610.8)	(99217.6)	(0.029)	(0.024)
- Bootstrapped p-value	0.012	0.006	0.000	0.000
Baseline controls	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer
MCPs	13	13	13	13
Observations	1001	1001	1001	1001
Mean of non-upgraded	1206519.7	1206519.7	0.665	0.665
SD of non-upgraded	1495124.7	1495124.7	0.472	0.472
Upgraded (ITT)	-430427.2**	-410492.9***	0.169**	0.146***
- Clustered SE	(104474.1)	(73391.5)	(0.036)	(0.033)
- Bootstrapped p-value	0.038	0.002	0.012	0.002
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	13	13	13	13
Observations	1001	1001	1001	1001
Mean of non-upgraded	1187541.9	1187541.9	0.669	0.669
SD of non-upgraded	1444847.7	1444847.7	0.471	0.471

Top panel shows comparison at farmer level, bottom panel shows comparison after weighting observations by milk quantity
Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

Table A4.9. Impact on Hygienic Quality (1st Upgraded MCP Excluded)

	Outcome variables			
	(1)	(2)	(3)	(4)
	TPC	TPC	TPC<1M	TPC<1M
Upgraded (ITT)	-520488.8***	-532998.8***	0.155***	0.164***
- Clustered SE	(144839.4)	(108085.3)	(0.044)	(0.031)
- Bootstrapped p-value	0.010	0.004	0.002	0.006
Baseline controls	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer
MCPs	12	12	12	12
Observations	596	596	596	596
Mean of non-upgraded	1231838.7	1231838.7	0.701	0.701
SD of non-upgraded	1600975.2	1600975.2	0.458	0.458
Upgraded (ITT)	-538323.1**	-508215.2***	0.147**	0.128*
- Clustered SE	(149091.3)	(81563.6)	(0.052)	(0.051)
- Bootstrapped p-value	0.014	0.002	0.028	0.066
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	12	12	12	12
Observations	596	596	596	596
Mean of non-upgraded	1201460.7	1201460.7	0.717	0.717
SD of non-upgraded	1579282.2	1579282.2	0.451	0.451

Top panel shows comparison at farmer level, bottom panel shows comparison after weighting observations by milk quantity

Standard errors clustered at MCP level in parentheses

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

4

Table A4.10. Impact on Hygienic Quality (MCPs Upgraded in 2018 Excluded)

	Outcome variables			
	(1)	(2)	(3)	(4)
	TPC	TPC	TPC<1M	TPC<1M
Upgraded (ITT)	-379228.3**	-395363.2**	0.108***	0.118***
- Clustered SE	(129298.8)	(120811.0)	(0.033)	(0.028)
- Bootstrapped p-value	0.028	0.032	0.000	0.006
Baseline controls	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer
MCPs	11	11	11	11
Observations	605	605	605	605
Mean of non-upgraded	1177818.6	1177818.6	0.715	0.715
SD of non-upgraded	1556762.5	1556762.5	0.452	0.452
Upgraded (ITT)	-369878.9**	-345947.0***	0.0914**	0.0865***
- Clustered SE	(119851.9)	(68529.7)	(0.0376)	(0.0209)
- Bootstrapped p-value	0.028	0.002	0.024	0.004
Baseline controls	No	Yes	No	Yes
Unit of analysis	Milk	Milk	Milk	Milk
MCPs	11	11	11	11
Observations	605	605	605	605
Mean of non-upgraded	1100453.2	1100453.2	0.745	0.745
SD of non-upgraded	1466924.6	1466924.6	0.437	0.437

Top panel shows comparison at farmer level, bottom panel shows comparison after weighting observations by milk quantity

Standard errors clustered at MCP level in parentheses

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

* p<0.10, ** p<0.05, *** p<0.01

5. The Competitiveness of Agricultural Markets: Evidence from an Aid for Trade Intervention in the Indonesian Dairy Value Chain

Mark Treurniet

Wide-spread market concentration in global value chains at the level of processors and exporters may cause non-competitive pricing and rent-capturing. I study an Aid for Trade intervention in the Indonesian dairy value chain that aimed to increase technology adoption among farmers and led to an increase in milk quality. Using a unique panel of quality and price data, I assess how the resulting quality rents are distributed between processor and producers. Farmers receive higher prices in the short run, but the processor captures the full quality rents after the intervention was completed.

5

5.1. Introduction

Fueled by disappointment about the (lack of) impact of conventional aid modalities, an increasing share of ODA nowadays focuses on creating opportunities for trade. According to OECD data, so-called Aid for Trade commitments have grown from an average of $22 billion in 2002-2005 to more than $50 billion in 2015 and 2016, and now amount to approximately 30% of total ODA (OECD/WTO 2017, OECD 2018). These Aid for Trade commitments are meant to help developing countries building "the trade capacity and infrastructure they need to benefit from trade opening" (WTO 2019) and are mostly focused on building economic infrastructure and productive capacity.

One of the top priorities of donor agencies in the Aid for Trade agenda is to connect suppliers from developing countries to modern value chains (OECD/WTO 2017). The additional value generated via such modern value chains is expected to raise smallholder income and create incentives for investments in production. Donor agencies often disperse development aid through the donor country's private sector. The private sector can leverage these donor subsidies by co-financing value chain interventions that help smallholder producers to adopt new technologies and increase the value of their produce. As a result of these interventions, the donor country's private sector may benefit from increased access to high-quality inputs and local consumer markets,[44] creating the promise of a win-win situation.

However, it is not evident that smallholder farmers stand to gain much from such interventions. This depends on the actual distribution of the value that is created among participants in the value chain. This insight dates back to Cochrane's (1958) classic "agricultural treadmill" theory. In an attempt to remain competitive, smallholders continuously have to adopt new technologies that reduce per unit costs of production. While early adopters will temporarily benefit from increased profit margins, under inelastic demand widespread adoption will push down prices. Profit margins dwindle and non-adopters may be forced out of business. Benefits, instead, mainly accrue to customers. This outcome is not driven by malevolent intent or behavior of any actor in the chain – it is driven by atomistic production combined with positive supply responses and steeply downward sloping demand.

---

[44] Though the term Aid for Trade might suggest otherwise, the produce is traded within value chains with international actors, but not necessarily exported.

This paper puts forward another mechanism that generates an agricultural treadmill. Global value chains for food commodities are characterized by widespread market concentration. Large players in the trade or processing sector have the ability to set prices, within limits, to maximize their own profits. Conditions may emerge where smallholders have to adopt new technologies in order to keep up with increasing private standards and create economic value, but fail to reap any of this value as additional income for themselves. Instead, market power enables downstream multinationals to extract the rent by paying lower prices. A variant of Cochrane's treadmill eventuates, where farmers have no choice but to adopt new technologies and fail to benefit from them.

While there is widespread suspicion in development circles that traders in agricultural markets in developing countries abuse their market power and exploit smallholders by offering "low prices", the available evidence, summarized and discussed by Dillon and Dambro (2017), suggests agricultural markets are rather competitive. However, available studies often focus on markets close to smallholder producers or to consumers. This allows studying large numbers of traders (Casaburi and Reed 2017) or local markets (Bergquist 2017). Little is known about the competitiveness of agricultural markets in developing countries dominated by a few large (international) firms because of significant barriers to entry. In an increasingly globalizing world the behavior of such large firms becomes increasingly important. For example, Casaburi and Reed (2017) study local cocoa markets in rural villages in Sierra Leone, concluding these are rather competitive. But lower down in that same value chain nearly all cocoa is purchased and traded by a single foreign firm – begging the issue of rent distribution between actors along the entire chain, rather than between smallholder producers and local traders.

The objective of this paper is to introduce a new approach to study the competitiveness of concentrated markets and present first evidence on how market imperfections may undermine the impact of development interventions aimed at 'upgrading' smallholder agricultural producers. Specifically, I study the distribution of rents associated with an Aid for Trade intervention in the Indonesian dairy value chain. The purpose of the intervention was to help smallholder cooperative members to increase the quality of milk produced, so that they will benefit from higher milk prices. Improved milk was sold to a large dairy processor. However, the distribution of quality rents between the cooperative and the processor critically depends on their relative market power. I present a model that predicts how milk prices respond, depending on assumptions with respect to distribution of market power.

The main result of this study is that the processor has significant market power. Although there was a delay in exercising that market power, the processor eventually captured the full intervention rents after the intervention was completed. This finding supports the notion of monopsonistic markets with only a very small share of the intervention's benefits trickling down to smallholder producers.

The remainder of this paper is organized as follows: Section 5.2 discusses existing literature on the competitiveness of agricultural markets, and Section 5.3 presents a simple model that relates the degree of market power to the distribution of rents arising from adopted technologies. The following sections discuss the context and intervention of my empirical study, as well as the data used to test prediction and the empirical strategy used. Using a unique panel of milk transactions between the cooperative and two processors, I analyze how the actual evolution of prices compares to model predictions. Section 5.8 concludes and discusses methodological lessons and implications for Aid for Trade interventions in value chains.

5.2. Competitiveness of agricultural markets

Whether agricultural markets are competitive is a long-standing and pertinent question (Dillon and Dambro 2017). The degree of competition has strong implications for the formation of prices and the distribution of rents along value chains. Moreover, it affects incentives for modernization and investment by smallholders, and shapes the effects of value chain interventions that are designed to help poor producers.

Against the background of common anecdotes and suspicions of non-competitive pricing by traders, Dillon and Dambro (2017) review existing evidence on the competitiveness of crop markets in Sub-Saharan Africa. Given the thinly available evidence on this important topic, Dillon and Dambro find that competitive markets are widespread and there is little empirical support for widespread rent-seeking by traders. Yet, the generalizability of these findings is limited by three key characteristics of the studies underlying their review.

First, where Dillon and Dambro mostly focus on markets for grains with a long shelf-life, the degree of competition in markets may be affected by the perishability of products traded. Perishability may increase transaction costs, which decreases the time and space dimensions of the market and, thus, the effective number of traders competing for the product. Using the idea that such constraints to competition can potentially be offset by any technology that decreases transaction costs, Muto and Yamano (2009) study the impact of phone coverage expansion in Uganda and indeed found that prices for perishable bananas increased, while prices for less perishable maize were unaffected. Though this may suggest that markets for bananas were not competitive before the phone coverage expansion, banana prices may also have risen as a direct consequence of decreased transaction costs.

Second, while most studies in Dillon and Dambro's review do not take into account quality differentiation, competition might be less intense for high-quality products, since more investments may be required for processing and marketing of high-quality products, and barriers to entry may thus be higher. A few studies do explicitly study local markets for high-quality products. Casaburi and Reed (2017), for example, randomly provided quality premiums to cacao traders in Sierra Leone and found limited differences in prices paid by treated and control traders, supporting the notion of competitive markets. Bernard et al. (2017) randomly provided famers with information about introduction of scales and quality labelling for onions on local markets in Senegal and found that this increased investments in quality, leading to 6-9% higher prices received by farmers. Cost-benefit analysis in the latter study shows that informed farmers increased their net revenues, suggesting that the market for high-quality onions was at least somewhat competitive.

Third, and potentially most seriously, while trader rent-seeking may not be widespread, more downstream processing and export firms may still be able to earn non-competitive rents. At this stage of the value chain, global markets are considerably concentrated with combined market shares for the four largest firms (CR4) estimated at 0.61 for cacao grinding (Gaji and Tsowou 2015), 0.41 for coffee export (Grabs 2017), and 0.42 for banana export (FAO 2014). Focusing on smaller geographical areas and considering product differentiation, many local markets are likely to be even more concentrated due to market segmentation. Studies reviewed by Dillon and Dambro, however, exclusively focus on the role of traders. Moreover, frequently used methodologies that rely on comparisons across a high number of traders or markets (Dillon and Dambro 2017, Casaburi and Reed 2017, Bergquist 2017), are not well-suited to proxy competitiveness of large markets that are dominated by a few processors or exporters. Given the great importance of the topic, it is important to avoid that preferred empirical methodologies dictate in which type of markets competition is researched (Ravallion 2012), so that the challenge to assess market power in large concentrated markets is entirely left to anecdotes. Yet, I am not aware of any study empirically testing whether processors or exporters pay competitive prices in developing countries.

This paper contributes to the empirical literature on the competitiveness of agricultural markets in developing countries by focusing on markets for *highly perishable* and *high-quality* milk where few large downstream *processors* control most demand.

5.3. Model

I model the distribution of intervention-induced quality rents between processors and producers. I distinguish three cases, each making different assumptions on market power and producing different predictions on the distribution of quality rents. The first case takes one extreme and assumes Bertrand competition at the side of the processors, so that the producer has full market power over the quality rent. The second case takes the other extreme and assumes that the processor acts as a monopsonist and has full market power (Swinnen and Vandeplas 2011, Sexton 2012). The third case bridges both extremes and assumes that the processor and producer bilaterally bargain over the quality rents (Collard-Wexler, Gowrisankaran, and Lee 2019).

5.3.1. General set-up and assumptions

Throughout this paper, I assume that processors are profit-maximizing and processors' profits equal the margin between the net value of processed products and the price paid for raw produce. As high-quality produce can be processed into higher value products, producers with the capacity to make these products (which will be referred to as high-quality processors) can make a higher profit on high-quality produce. More formally, the high-quality processors' profit per unit produce π is increasing in hygienic quality q and decreasing in the price paid to the producer p:

$$\pi = \alpha + \beta q - p, \qquad (5.1)$$

where $\alpha + \beta q$ is the value of milk and $\beta > 0$.

The producer is assumed to maximize its financial profit. The producer's marginal utility per unit produce is thus increasing in the price, but decreasing in the marginal cost of production:

$$u = p - c, \qquad (5.2)$$

The intervention comes at a one-time cost K, which will be sunk after implementation of the intervention, and raises quality from q_0 to q_1 and marginal costs of production from c_0 to c_1, where $\beta(q_1 - q_0) > (c_1 - c_0)$. The quality rents per unit produce then equal:

$$R \equiv \beta(q_1 - q_0) - (c_1 - c_0), \qquad (5.3)$$

I finally assume that there is sufficient competition for low quality produce, so producers can sell produce of any quality for a fixed price of:

$$p_0 \equiv \alpha + \beta q_0, \qquad (5.4)$$

5.3.2. Case 1: Bertrand competition

Let there be multiple high-quality processors who compete against each other in Bertrand price competition.

Processors that offer a price below the value of produce will not find sellers, as they will be outbid by other processors. Processors that offer higher prices make losses and will eventually exit the market. In equilibrium, therefore, prices equal the value of produce:

$$p_B = \alpha + \beta q, \qquad (5.5)$$

This leads to a post-intervention price of:

$$p_B = \alpha + \beta q_1 = p_0 + \beta(q_1 - q_0) = p_0 + (c_1 - c_0) + R, \qquad (5.6)$$

When the intervention increases quality from q_0 to q_1, the basic price remains constant, while the price increases by $\beta(q_1 - q_0)$. The high-quality processor does not make a profit and the producers capture the full intervention rents.

5.3.3. Case 2: Monopsony

Let there be one profit-maximizing high-quality processor who has monopsonistic power over the quality rents.

In line with the empirical observation that most changes in the price function over time occur in the basic price, assume that the high-quality processor offers a linear price function to the producer:

$$p_M = a_M + \beta q, \qquad (5.7)$$

where the basic price a_M is controlled by the high-quality processor.

Since the processor has full market power over the quality rent, it can set the producers at their outside option and extract all rents by setting a basic price lower than the outside option:

$$a_M = \alpha - R, \qquad (5.8)$$

This leads to a post-intervention price of:

$$p_M = p_0 + (c_1 - c_0), \qquad (5.9)$$

When the intervention increases quality from q_0 to q_1, the basic price decreases by R, while the price for high-quality produce increases by $(c_1 - c_0)$. The high-quality processor captures the full intervention rents and the producers are set at their outside option.

5.3.4. Case 3: Nash bargaining

Let there be one high-quality processor who bargains with the cooperative over the intervention rents.

In line with the empirical observation that most changes in the price function over time occur in the basic price, assume that the high-quality processor offers a linear price function to the producer:

$$p_N = a_N + \beta q, \qquad (5.10)$$

where a_N is the result of bargaining.

The Nash Bargaining solution is the solution to:

$$\begin{aligned}\max_{\alpha_N}[\pi - 0]^{\gamma}[u - u_0]^{1-\gamma} &= [\alpha + \beta q_1 - p_N]^{\gamma}[(p_N - c_!) - (p_0 - c_0)]^{1-\gamma} \\ &= [\alpha - a_N]^{\gamma}[R - (\alpha - a_N)]^{1-\gamma}, \qquad (5.11)\end{aligned}$$

where $u_0 \equiv p_0 - c_0$ and $\gamma \in (0,1)$ parametrizes the relative bargaining power of the high-quality processor.

Solving the above optimization problem yields:

$$a_N = \alpha - \gamma R, \qquad (5.12)$$

This leads to a post-intervention price of:

$$p_N = p_0 + (c_1 - c_0) + (1 - \gamma)R = \gamma p_B + (1 - \gamma)p_M, \qquad (5.13)$$

When the intervention increases quality from q_0 to q_1, the basic price decreases by γR, while the price increases by $(c_1 - c_0) + (1 - \gamma)R$ and becomes a weighted average of the prices predicted by the first two cases. The high-quality processor captures a share γ of the intervention rents and the producers receive a share $(1 - \gamma)$.

5.3.5. Graphical representation

Figure 5.1 graphically depicts these dynamics. Quality delivered by the producer q increases over the x-axis and the price paid by the processor moves along the y-axis. Initially, the quality supplied by the producer was relatively low as in q_0. The processor should compensate at least the outside option p_0 and therefore the price function goes through point (q_0, p_0). Due to the intervention, the quality delivered by the producer increased from q_0 to q_1, as illustrated by arrow (1). Under Bertrand competition assumptions, the processor would now pay net price p_B, which exceeds p_0. However, in the Monopsony case, the processor can again set the producers at their outside option by shifting the price function downward, as illustrated by arrow (2). Producers will now only be compensated for the increase in their marginal costs $(c_1 - c_0)$. Finally, in the Nash bargaining case, the downward shift in the price function is the result of a bargaining process. The downward shift is increasing in the relative bargaining power of the processor, but smaller than the shift in the Monopsony case.

5

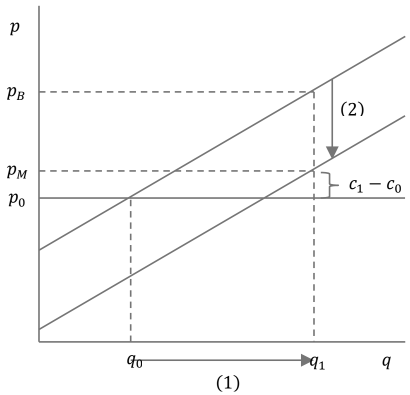

Figure 5.1. Graphical representation of special case

5.3.6. Timing of rent-capturing

If the high-quality processor has substantial market power, it may have strategic motives not to decrease the price function immediately. First, technology adoption takes time. As long as the producer is increasing its investments in technology adoption, the processor has an incentive to keep prices high to motivate the producer to continue this. Second, in a long-lasting buyer-seller relationship it may be difficult to decrease nominal prices. An alternative way to capture rents would be to wait until the nominal outside option increases, for example due to a correction for inflation. The processor then captures rents by not increasing its price as much. These considerations imply that short-run rent distributions may not reflect the distribution of market power, but predict that the share of the rent captured by the processor will gradually increase.

5.4. Empirical context and intervention

The Indonesian dairy value chain is characterized by many smallholder producers selling their milk to local dairy cooperatives by bringing it to their Milk Collection Points (MCPs). The dairy cooperatives sell most of their milk to a small number of dairy processors, which together control most of the local raw milk market. Raw milk is then processed into consumer products like liquid milk, sweet condensed milk, powdered milk, yoghurt and ice cream, and marketed to consumers (Wouters 2009, Morey 2011). Indonesia is a net importer of (high-quality) dairy products, so world market prices essentially determine the maximum price that processors can charge to domestic consumers (Wouters 2009, Morey 2011, Indonesian Ministry of Trade 2019b, a).[45]

Producing milk with good hygienic quality[46] is an ongoing challenge in the Indonesian value chain. Although heating can kill bacteria in milk, bacterial contamination is positively associated with the presence of toxins that will remain after heating. Milk with good hygienic quality is therefore better suited for mixing with imported milk powder.[47] Furthermore, it can

[45] Government regulations that required dairy processors to mix locally produced milk with imported milk power were abolished in 1998 as part of the IMF readjustment program (Wouters 2009).

[46] The term hygienic quality is used for low bacterial contamination, as measured by the TPC. This is to distinguish from other quality dimensions, like the chemical milk composition, as measured by the Total Solids content.

[47] When one aims to comply with desired standards for the final product, one can use more fresh milk if the hygienic quality is higher.

be processed into higher value products like cheese and infant milk. While the Indonesian National Standard (SNI) prescribes that the Total Plate Count (TPC) should not exceed 1M colony-forming units per milliliter (cfu/mL), actual TPCs may structurally exceed this maximum for part of the raw milk produced.

Proper behavior is seen as the key to improve hygienic quality. Producers should, for example, clean and dry the cow's teats before milking, throw away the first milk, which contains most bacteria, filter the milk adequately, and use a proper and clean milk can.

To incentivize the local dairy cooperatives to produce hygienic milk, a multinational processor offers local cooperatives a higher price for better quality milk. Bonuses are given if the TPC is lower than 1M, 500k and 100k cfu/mL. These bonuses are substantially higher than the bonuses for hygienic quality that other dairy processors offer to the dairy cooperative in this study, who at most give a penalty if the TPC exceeds 5M cfu/mL. Other price determinants are the Total Solids content as a measure for milk composition, the Freezing Point as a measure for adulteration of milk and the presence of Antibiotics.

Mid-2013, a public-private partnership (PPP) consisting of this multinational dairy processor, two local dairy cooperatives and several supporting non-profit organizations was granted a subsidy of several million euros to help the two local cooperatives and their members to, among other things, increase the hygienic quality of their milk. The project plan, as approved by the national donor, explicitly stated "increased profit and financial sustainability" for the members of the local cooperatives among the main results of this project.

As part of the intervention, the partnership supported one of the two local cooperatives[48] to upgrade MCP facilities, train its associated farmers, and measure the quality delivered by farmers individually, so that incentives can be implemented accordingly.[49] This package of interventions (which will be referred to as the intervention) was consecutively rolled out to seven MCPs. As a pilot, the intervention at the first MCP was fully paid by the PPP subsidy and the multinational processor, but the cooperative paid 75% of the investment costs of the

[48] Other interventions were implemented at the other cooperative, which remains out of the scope of this study.

[49] Before the upgrade, quality tests were performed at group level. After the upgrade, (i) quality was tested at individual level, and (ii) hygienic quality was measured more precisely, which allowed (iii) new bonuses to be provided for high quality milk.

six following MCPs. After upgrading seven MCPs, the cooperative decided to stop co-financing and expanding the intervention to other MCPs.

The data collected for this paper indicates that the hygienic quality improved over the course of the intervention. Chapter 4 compared quality across individual farmers from upgraded and non-upgraded MCPs and attributed a significant part of the increase in hygienic quality to the intervention. Impacts were suggested to have decreased over time, but remained significant. Chapter 4 further showed that, as a side-effect, the intervention also increased the Total Solids content of milk. Although the value created by the increase in Total Solids contents is relatively small, I do correct for this in my analysis.

Since hygienic quality increased, producers' incomes increase *if* the price function has not shifted. However, if increased quality causes the price function to shift downward, the processor captures a larger share of the rents and producers may not benefit as much as suggested by the project plan.

5.5. Data

5.5.1. Panel of transactions

To empirically test model predictions, I used a six-year panel of quality measures of milk delivered and prices paid by the high-quality processor and the other main buyer of the cooperative. Together these processors buy roughly 85% of the milk produced by the cooperative. The panel starts in January 2013, which is half a year before the subsidy was granted to the PPP. The panel runs until December 2018, which is more than three-and-a-half years after the start of project interventions and nine months after the cooperative decided to stop co-financing and expanding the intervention to other MCPs. For 2013, I have monthly aggregated quality measures and prices,[50] while for the following years I have milk truck-level quality and prices.

[50] For part of 2013, quality measures of milk delivered to the other processor are incomplete.

5.5.2. Price functions

This data implicitly contains the price functions used by the processors. The high-quality processor offers incentives for (i) the Total Solids content as an indicator for milk composition, (ii) the Total Plate Count (TPC) as an indicator for hygienic quality, (iii) the Freezing Point as an indicator for adulteration, and (iv) the presence of Antibiotics. First, the bonus for Total Solids is linear and similar in magnitude to the bonus set by the other processor. Second, bonuses for TPC are given if TPC is smaller than some cut-off points. The TPC bonuses are much larger in magnitude than those set by the other processor and associated to lower cut-off points. Further, the bonus for the Freezing Point was applied until 2015, relatively small in magnitude and only triggered for less than 2% of milk deliveries, while penalties for Antibiotics were severe, but only triggered for less than 0.1% of milk deliveries,[51] so these quality measures explain a negligible part of the variance in prices.

5.6. Empirical strategy

To test my model predictions, I first controlled for other variation in prices. I subsequently constructed proxies for the prices predicted by the Bertrand and the Monopsony cases of the model, so that I could compare the actual high-quality processor price evolution with these predictions.

5.6.1. Correcting for other variation

As follows from the price functions, differences between the prices set by both processors may be explained by differences in Total Solids, for which trends by processor over time are shown in Figure 5.2. Since the high-quality processor receives relatively more afternoon milk, the Total Solids content is higher in milk delivered to the high-quality processor than in milk delivered to the other processor. The intervention studied in this paper further contributed to the differences in Total Solids content (Chapter 4). To control for differences in Total Solids contents, I calculated what the price of the high-quality processor would have been according to its own price function if the Total Solids content would have been equal to the Total Solids content delivered to the other processor. In other words, I calculated $p_H(TPC_H, TS_O)$, where

[51] As frequencies are based on truck-level quality indicators, they are missing for 2013. The penalty for Antibiotics was triggered once in February, October and December 2014, and February 2015.

$p_H(\cdot)$ is the price function of the high-quality processor, TPC_H is the TPC of the milk delivered to the high-quality processor, and TS_O is the average Total Solids content of milk delivered to the other processor in this month.

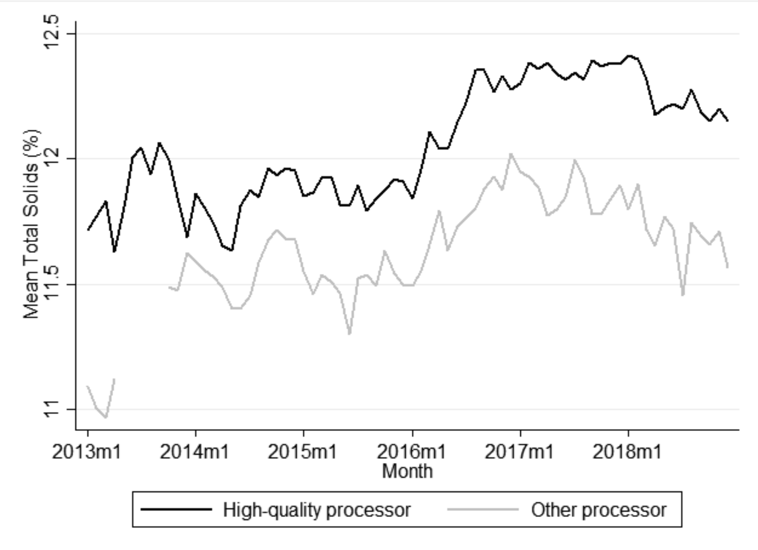

Figure 5.2. Mean Total Solids contents by processor

5

Further, time-invariant differences, like agreements about the transport of milk, may explain structural differences between the prices set by the two processors. I controlled for this by subtracting from the high-quality processor's price the mean difference between the high-quality processor's price and the other processor's price during January-April 2013, the months before the subsidy was rewarded for which I have data on Total Solids content and prices for both processors.

Finally, to protect confidential price information, I only show prices after correcting for these time-invariant differences and converting them to indexes, where the price of the other processor in January 2013 was set to 100.

5.6.2. Constructing proxies for model predictions

Figure 5.3 shows mean TPC over time for milk delivered to both processors, while vertical lines mark specific events related to the intervention. The first vertical line marks the decision on granting the subsidy to the PPP, which was taken on June 28, 2013. While some data is missing for the other processor for 2013, the mean TPC of milk delivered to both processors structurally exceeded the national standard of 1M cfu/mL until the opening of the first MCP in June 2015, as marked by the second vertical line. Starting from this moment, the mean TPC delivered to the high-quality processor started to decreases over time, mainly because the proportion of milk coming from upgraded MCPs increased. From August 2016, marked by the third vertical line, the high-quality processor only received milk from upgraded MCPs with a TPC below 1M cfu/mL.[52] Although the mean TPC increased again as more time passed after the intervention trainings, it remained below 650k cfu/mL during the whole post-intervention period studied. Meanwhile, the mean TPC delivered to the other processor also decreased, but remained well above 1M cfu/mL. The fourth vertical line marks the decision of the cooperative to stop co-financing and expanding the intervention to other MCPs, which was made in March 2018.

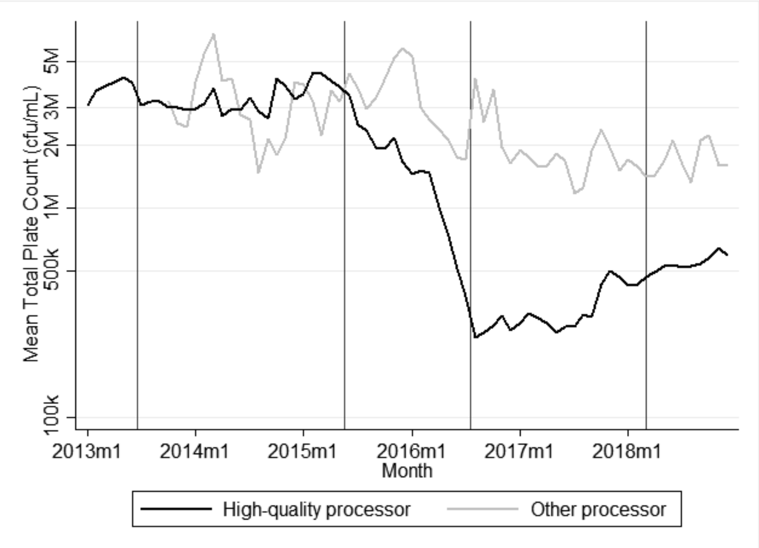

Figure 5.3. Mean Total Plate Count by processor

Motivated by equation (5.6), I constructed the Bertrand competition prediction as the sum of the price set by the other processor and the quality premium. I calculated the quality premium as the bonus according to the high-quality processor's price function relative to some reference quality category. I base this reference category on the mean quality delivered to the other

[52] From August 2016, the high-quality processor rejected milk if the TPC exceeds 1M cfu/mL. Since in practice milk was very rarely rejected, I did not explicitly study rejection in this paper.

5

processor. As Figure 5.3 shows that while the quality of the milk delivered to the high-quality processor decreased to below 1M cfu/mL, the TPC of milk delivered to the other processor mainly remained between 1M and 2M cfu/mL, this latter category is therefore chosen as reference category.[53]

Following equation (5.9), I constructed the Monopsony prediction as the sum of the price set by the other processor and the increase in marginal costs. Most of the costs of producing higher quality milk are part of the fixed investment cost K and thus do not increase with milk production. However, since the intervention included the introduction of more precise testing of individual milk samples, the intervention also increased variable costs. While individual milk samples are tested by the high-quality processor, the associated costs are subtracted from total milk payments to the cooperative. I therefore used testing costs per kg milk as a proxy for the increase in marginal costs.[54] As increased variable effort exerted to implement hygienic practices is not quantified and not included in the construction of the Monopsony prediction, the Monopsony prediction is a conservative estimate of the true Monopsony prediction.

Following equation (5.13), the Nash bargaining prediction is a weighted average of the Bertrand competition prediction and the Monopsony prediction. The weights depend on the unknown relative bargaining power of the cooperative and the high-quality processor, which were estimated by this study.

5.7. Results

Actual prices over time and proxies for my model predictions are shown in Figure 5.4, which includes the same vertical lines as Figure 5.3. I compared the actual price paid by the high-quality processor with two benchmarks: the Bertrand competition price and the Monopsony price. The Bertrand competition price, at the one extreme, is calculated as the other processor's price plus the high-quality processor's quality premium, and is assumed to be paid by the high-quality processor if it has zero market power over the intervention-induced quality rents. The Monopsony price, at the other extreme, equals the other processor's price plus the additional

[53] As from August 2016, no milk with a TPC above 1M cfu/mL was received by the high-quality processor, the quality premiums in this period could not be calculated from the transaction data. Since incentives for TPC levels rarely changed, quality premiums were assumed to remain unchanged after July 2016.

[54] Costs of the simpler tests that are applied at non-upgraded MCPs are negligible.

marginal costs, and represents the minimum price that the high-quality processor has to pay if it has full monopsonistic power. Following equations (5.6) and (5.9), the difference between the Bertrand competition price and the Monopsony price represents the intervention-induced quality rent. By comparing the actual price paid by the high-quality processor with these two extreme predictions, I estimated the share of the quality rents captured by the high-quality processor. Following equation (5.13), this share can be interpreted as an estimate for the relative market power that has been exercised by the high-quality processor.

The first vertical line marks the decision on granting the subsidy to the PPP. As noted before, I have shifted the price of the high-quality processor to equal the price of the other processor before this date. The following results can thus be interpreted as Difference-in-Differences, and its causal interpretation relies on two critical assumptions: (i) While the cooperative only supplies a small part of total milk in the market, and investments in hygienic quality were unlikely to be broadly replicated at other cooperatives during the period studied, I assume that the intervention did not affect general market prices for high- and low-quality milk. (ii) While consumer demand for high-quality milk is increasing, and the high-quality dairy processor had chosen to invest in improving the hygienic quality of milk, I assume that the difference in the value of high-quality and low-quality did not (entirely) disappear during the study period.

The second vertical line marks the opening of the first upgraded MCP. In the months before the opening of the first MCP, the price of the high-quality processor falls short of the other processors price. I speculate that the cooperative might have temporarily accepted this lower price, as it also received investment in the form of the fully subsidized intervention at the first MCP.

From the opening of the first upgraded MCP, the Bertrand competition price increased relative to the Monopsony price as the proportion of milk received from upgraded MCPs increased. The difference between these prices represents the increasing quality rents. The actual price paid by the high-quality processor also increased.

From the third vertical line onwards, the high-quality processor only received milk from upgraded MCPs. The high-quality processor price closely followed the price as predicted by the Bertrand competition case of the model. Between August 2016 and February 2018, the period between the third and fourth vertical line, rent-capturing by the high-quality processor was estimated at $\hat{\gamma} = 0.14$, which suggests that the cooperative received a return to investments in quality in the short run.

The fourth vertical line marks the decision of the cooperative to stop co-financing and expanding the intervention to additional MCPs. After this point, the price of the other processor also increased, reflecting rising import prices and increasing local demand.[55] The Bertrand price increased accordingly, but the price offered by the high-quality processor's price did not match this overall increase. Instead, it converged towards the Monopsony price, offering the cooperative nothing more than its reservation price – so that the cooperative is indifferent between offering its (high-quality) milk to the high-quality and low-quality processor. In the process, quality rents have shifted from the cooperative and its members to the processor. Rent-capturing by the high-quality processor was estimated to have increased to $\hat{\gamma} = 1.04$ for the period July-December 2018, which suggests that the cooperative does not receive any return to its investment in the long run.

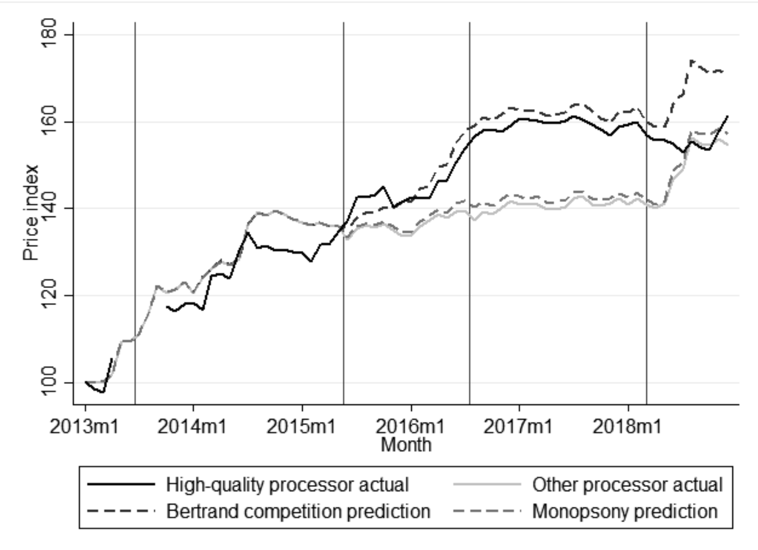

Figure 5.4. Prices relative to model predictions

To compare aggregate benefits and costs to the cooperative, I multiplied the difference between the high-quality processor's price and the Monopsony price with the quantity delivered to the high-quality processor, and summed over the years 2014-2018.[56] The aggregate returns were

[55] Local demand could have increased due to higher consumer demand and political pressure on milk processors to increase local milk sourcing.

[56] Quantity was missing for 2013.

estimated to be 89% of the investment costs borne by the cooperative, which suggests that the short-run benefits were insufficient to cover the cooperative's full investment costs.

5.8. Discussion

This paper revamps Cochrane's (1958) agricultural treadmill hypothesis, and suggests that analogous dynamics may eventuate when processors or exporters have substantial market power in global value chains. While such companies may require smallholder farmers to continuously adopt new technologies and improve quality, they can extract the bulk of the rents by adjusting the prices they offer to the inputs they purchase. In the process, non-adopting smallholder producers lose out and see their income deteriorate as they face a lower price for the (low-quality) output that they produce.

As a case study, I studied the competitiveness of the local market for highly perishable and high-quality milk in Indonesia where demand is largely controlled by a few large downstream processors by studying the price effects of an Aid for Trade intervention that increased the hygienic quality of milk. In the short run, prices seemed to be in line with fairly competitive pricing. Yet, after the cooperative had decided to stop co-financing and expanding the intervention to other MCPs, an increase in the other processor's price was not followed by an increase in the high-quality processor's price. As a result, quality rents started to shift from the cooperative to the processor. Towards the end of our study period the processor captured the full quality rent, suggesting it has monopsonistic power.

5

Two factors may explain why the high-quality processor did not capture full quality rents earlier. First, as long as the cooperative considered co-financed the intervention, the processor had an incentive to keep prices high to motivate the cooperative to continue upgrading other MCPs. Second, the price predicted by the Monopsony model remained rather stable in the period after the introduction of the intervention. This implies that in order to capture rents earlier, the high-quality processor would have needed to decrease nominal prices, which may be relatively difficult in a long-lasting buyer-seller relationship.

An alternative explanation for the observed agricultural treadmill could be that increased supply of high-quality milk and decreased supply of low-quality milk might have caused the quality premium to decrease. Although I do not have direct data on market-wide developments in hygienic quality, several observations make this alternative explanation unlikely: First, the cooperative in this study supplies only a small portion of total milk in the market. Second, the high-quality processor was the only dairy processor that offered substantial incentives for high levels of hygienic quality. As significant investments were needed to increase hygienic quality and were enabled by an external subsidy, these investments were unlikely to be broadly replicated during the period studied. Third, if intervention spillovers would have affected quality, they would most likely be largest within the cooperative. Although Figure 5.2 does show that hygienic quality of milk delivered to the other processor increased over time, spillovers were insufficient to meet the Indonesian National Standard of 1M cfu/mL. Fourth, the price increase that made the quality premium disappear was likely the result of increased demand rather than changes in supply.

If the market would further develop in the coming years, and further increase quality standards, then farmers that do not manage to increase their quality could struggle to deliver to modern value chains and, finally, be driven out of the market. Although the cooperative did not benefit from long-term increased prices, it may therefore be better off in the long run compared to other cooperatives that did not manage to increase quality. This would again be in line with Cochrane's theory, which predicts that non-adopters may be forced out of business as a result of increased competition among producers.

While studying price effects after an exogenous increase in quality allowed me to proxy the competitiveness of the local market for high-quality raw milk in Indonesia, some methodological lessons can be learned from this study. My empirical analysis suggest that identification of rent-seeking critically relies on four important factors:

First, rent-seeking may only be identified at aggregate level. If one price matrix is set for milk from a group of farmers, then within this group prices for high quality milk may be higher than prices for low quality milk. When an intervention now helps part of the farmers within this group to increase their quality, the single price matrix used may cause intervention farmers to receive higher prices than non-intervention farmers. However, rent extraction may be effectuated by slowly decreasing the basic price, and thus lowering prices for both high and low quality milk. Within-in group differences then do not reflect the causal effect of quality improvement on prices, and conceal rent extraction at the aggregate level.

Second, prices should be studied over a sufficiently long period, as (i) the actor with market power may have strategic motivations to postpone capturing a larger share of rents, and (ii) in long-lasting buyer-seller relations, it may be difficult for this actor to alter nominal prices in his benefit. Studying an insufficiently long period may result in an overestimate of the competitiveness of markets.

Third, the availability of a credible proxy for the outside option is crucial to identify whether quality rents are shifted to another actor, since (i) intervention rents may not be shifted instantly, (ii) other factors may also influence market prices. If such a credible proxy for the outside option is not available, an unknown bias might be introduced.

Fourth, intervention rents should be sufficiently large relative to other factors that may cause some variation around the parallel trends in prices. If intervention rents are too small, the proxy for the competitiveness of the market becomes too rough.

The results of this study have strong implications for Aid for Trade interventions in value chains. Policy makers should not simply assume that smallholder producers benefit from interventions that improve the value of their produce. Instead, market conditions shape whether smallholder benefits from such interventions will sustain over time. Monopsonistic power may limit or completely eliminate smallholder benefits, especially if interventions lead to increased market segmentation. In such cases, Aid for Trade interventions risk increasing the speed of the agricultural treadmill, rather than supporting smallholder incomes.

If markets are not competitive, an alternative approach to enhance smallholder producers' welfare could be to improve their outside option. For example, the cooperative in this study co-owns a small processing factory. However, its output is constrained by the marketing capacity of the cooperative. Supporting the cooperative's marketing capacity might increase cooperative income as well as increase the outside option for raw milk. According to the model's logic, this would enable the cooperative to negotiate higher prices for raw milk.

To effectively design Aid for Trade policies to benefit smallholder producers in the future, we need to continue to improve our understanding about how market function. Donor agencies may play a crucial role in ensuring that researchers get access to important price data, so that the impact of PPPs on smallholder farmers can be properly assessed under different market conditions.

5.9. Epilogue

The analysis in this Chapter is based on quality-adjusted prices paid by the dairy processor to the farmer cooperative between January 2013 and December 2018. The results were shared with the dairy processor and the cooperative in March and April 2019. In February 2020, I received an email from the high-quality processor containing a graph with more recent prices, extending until the end of 2019. While the price data of these two time series are not directly comparable (there was no quality adjustment in the latter price time series, and model predictions were not included), it appears as if the high-quality processor now pays prices exceeding those paid by the other dairy processor. In the meantime, the cooperative has also increased its investments in quality. If I receive more price data conform request, I will incorporate them in my analysis and report the new results in a new publication.

6. Do Public-Private Partnerships in Agricultural Value Chains Reach Poor Farmers? Evidence from the Indonesian Dairy Industry

Mark Treurniet, Jos Bijman, Erwin Bulte and Marlene Roefs

Public-private partnerships (PPP) are an increasingly important component of the Aid for Trade development strategy, aiming to transform agro-food value chains in developing countries to promote integration of smallholders into high-value commodity markets. But which farmers are included in such initiatives? We study which farmers are reached by a PPP that is led by a multinational dairy firm and working in the Indonesian dairy sector, and find that the PPP intervention (i) reaches smallholder producers that are relatively wealthy, and (ii) increases economic inequality among the population of smallholders (all members of a dairy cooperative). We argue that the PPP intervention may have created "winners" and "losers", and that non-included poor smallholders may be worse off as a result of lower prices for their output.

6.1. Introduction

One of the global challenges today is agricultural development in low-income countries. The majority of the world's poor reside in rural areas, and their economic fate depends on the performance of the agricultural sector (e.g. Byerlee et al. 2008, Dercon and Gollin 2014). Studies have shown that economic growth in agriculture is the most effective way to lift people out of poverty. It has large multiplier effects in early stages of economic development (Haggbladde, Hazell, and Dorosh 2007), and income growth originating in agriculture raises income of the poor much more than growth originating elsewhere in the economy (Ligon and Sadoulet 2007, Christiaensen, Demery, and Kuhl 2011). Moreover, due to various forward and backward linkages, agriculture could be an important sector promoting economic modernization. Finally, raising smallholder productivity helps to attenuate concerns about how to feed a growing world population. Unfortunately there is little consensus on how to promote agricultural development in low-income countries.

The liberalization of international trade has not delivered on its promise to modernize and intensify smallholder farming in low-income countries. Trade liberalization has by-and-large failed to transform rural areas in developing countries into productive regions offering gainful opportunities for smallholder farmers. Important causes of the inability of poor farmers to produce for export markets are non-tariff barriers, weak infrastructure, and various supply constraints (e.g. Stiglitz and Charlton 2013). This realization invited launching of the 2005 'Aid for Trade' initiative, aiming to "help developing countries, particularly LDCs, to build the supply-side capacity and trade-related infrastructure that they need to assist them to implement and benefit from WTO Agreements and more broadly to expand their trade" (WTO Hong Kong Ministerial Declaration 2005).

To implement the Aid for Trade agenda, the international development community has up-scaled its investments in productive sectors. The share of Aid for Trade investments in total ODA has increased from an average of $22 billion in 2002-2005 to more than $50 billion in 2015 and 2016, and now amounts to approximately 30% of total ODA (OECD/WTO 2017, OECD 2018). While the bulk of these investments aims to build economic infrastructure, almost two-fifth of Aid for Trade funding seeks to build productive capacity. Most of the latter funding is directed towards developing and supporting agro-food value chains – enabling farmers to produce more and better crops, and improving linkages of smallholders to domestic and foreign markets. Rather than focusing on reforming macro-, trade- or price policies, the development focus has shifted towards exploring institutional innovations to reduce inefficiencies in value chains (Barrett et al. 2012).

Aid for Trade interventions in food value chains often involve the private sector via public-private partnerships (PPPs).[57] Donors have multiple reasons to involve the private sector, including expected gains from the expertise, network and efficiency of the private sector, as well as the promise of leveraging public investments by private co-funding (Poulton and Macartney 2012). On the other hand, the private sector's profit motive may, depending on the context, not necessarily align well with development objectives. Private firms are likely drawn to partnerships by the prospect of enhanced access to products and local consumer markets, and may seek to reduce transaction costs by targeting relatively efficient and 'large' producers. Such producers are unlikely to be the poorest of the poor, so the private firm's efficiency considerations may conflict with distributional and anti-poverty concerns that public agencies generally have. This trade-off mirrors the well-known debate about "outreach" versus "financial sustainability" in the world of microfinance (e.g. D'Espallier et al. 2017, Mia and Lee 2017). Although there is an extensive and inconclusive discourse on whether and how the bottom of the pyramid are (to be) included in the value chains that are being supported by the PPP interventions (Bitzer and Glasbergen 2015), little is known about the actual targeting strategies of firms in PPPs.

In this paper we analyze the inclusiveness of a PPP led by a multinational firm, and ask whether the PPP intervention reaches a non-random subsample of relatively privileged producers as preferred local partners. Specifically, we consider the case of a dairy processor contracting with a local cooperative in Indonesia, assisting (and incentivizing) small-scale dairy farmers to produce high-quality milk for processing and subsequent sale in local retail markets. The intervention involves upgrading local 'Milk Collection Points' (MCPs), and training associated dairy farmers to produce high-quality output. We consider which farmers are associated to the MCPs that are selected to benefit from this intervention. We also analyze the intra-cooperative distributional consequences of the intervention, and compare economic outcomes for treated

[57] A PPP is "a long-term contract between a private party and a government entity, for providing a public asset or service, in which the private party bears significant risk and management responsibility, and remuneration is linked to performance" (World Bank 2019b).

and non-treated dairy farmers within the same dairy cooperative – exploring whether there are winners and losers.

We find that the PPP intervention reaches local dairy farmers that, compared to their peers, are wealthier, and own more cows. We also demonstrate that the economic fates of the treated and non-treated farmers are closely linked because milk payments to all producers are based on one and the same price matrix (linking per unit milk prices to the quality of milk that is delivered). In the case we consider, treated farmers are made better off and, apart from potential positive spillover effects, non-treated farmers are made worse off as a result of the intervention. In light of the observation that non-treated farmers predominantly belong to an under-privileged group of local producers and that the intervention was funded with public money, our results raise an important dilemma for policy makers.

We are among the first to consider the inclusiveness of PPPs led by multinational firms involving smallholders in agricultural value chains, and explore the distributional consequences. The focus is on intra-cooperative consequences as all subjects in our sample are member of the same cooperative.[58] Due to the increasingly prominent role of PPPs in international aid strategies this is now an important issue, complementing earlier work on, say, membership selection in local cooperatives (e.g. Bernard and Spielman 2009, Fischer and Qaim 2012) and outsourcing strategies of multinational firms engaged in contract farming (Barrett et al. 2012).

The remainder of this paper is as follows. In Section 6.2 we discuss the case study and the nature of the PPP intervention. Section 6.3 introduces our data. In Sections 6.4 and 6.5 we present empirical results focusing, respectively, on inclusiveness and distributional outcomes. The discussion ensues.

6.2. Study setting

We study the inclusiveness and distributional effects of an Aid for Trade intervention in the Indonesian dairy value chain. We consider outcomes for the members of a large dairy cooperative on Java. About 3,700 dairy farmers deliver almost all their milk to one of the cooperative's 31 Milk Collection Points (MCPs). The cooperative aggregates the milk and sells

[58] See Bouma and Berkhout (2015) for another perspective, focusing on selection on where PPPs are initiated.

most of it to dairy processing companies. The cooperative also provides to its members (veterinary and production) inputs, extension services, financial services and access to a health facility.

In an effort to, among other things, increase the compositional and hygienic quality of milk, promoting adherence to Indonesian National Standard (SNI),[59] a public-private partnership (PPP) was initiated between the cooperative, one of its main buyers, and several supporting not-for-profit organizations. The PPP received a subsidy of several million euros by a European donor country. The implementation strategy included several steps, discussed below, and was "rolled out" during the period 2015-2018 in seven MCPs out of the full set of 31 MCPs. In this paper, we are interested in the outcome of the selection process, and study the actual inclusiveness and distributional effects of this intervention. The analysis below is therefore based on a comparison of farmers from selected and non-selected MCPs. To analyze the inclusiveness issue, these farmers are compared using baseline data. To probe the distributional issues associated with the intervention, we use follow-up administrative data from the cooperative.

The PPP implementation strategy involved provision of material inputs and a training component, complemented by a subtle institutional reform. First, selected MCP facilities were upgraded. This implied installing a new registration and sampling system, and building new cleaning facilities for milk cans. Farmers supplying to upgraded MCPs received loans to buy new milk cans, and complementary filters and buckets. Second, four to eight months before opening upgraded MCPs, farmers were invited to attend a socialization and information meeting. Farmers also attended four to six two-hour trainings on hygienic practices,[60] and were visited by extension officers using checklists to monitor practices. Individual milk samples were taken to monitor hygienic quality, providing additional information for tailor-made feedback. Third, after upgraded MCPs opened, farmers received prices based on the quality of the milk they supplied individually, as opposed to prices based on the average quality of milk

[59] Compositional quality is based on solid content (fat and proteins). According to Indonesian National Standard, there should be at most 1M bacterial colony-forming units per milliliter of milk (cfu/mL).

[60] This includes information on: cleaning and drying the cow's teats before milking, throwing away the first milk, filtering the milk adequately, and using a proper and clean milk can.

produced by all farmers in a so-called local payment group.[61] To enable accurate measurement and payment, hygienic quality was more precisely measured with new equipment.

Chapter 4 evaluates the impact of the PPP intervention, and demonstrates that the intervention improved the hygienic and compositional quality of milk delivered to the cooperative. As a result, treated farmers appear better off. But who are these treated farmers, and how do their outcomes affect economic returns of their colleagues supplying to non-upgraded MCPs?

6.3. Data

MCP-level baseline variables are constructed from cooperative administrative data, including the number of farmers delivering milk, and quantities of milk supplied. We also have a variable indicating whether any cooperative board members delivered to the MCPs. At the farmer-level, individuals are labelled as treated or non-treated, depending on the upgrading status of the MCP to which they delivered their milk in 2014 (virtually no switching of MCPs takes place in response to the intervention or otherwise, as MCPs are dispersed and transport of fresh milk is cumbersome). We refer to this treatment status as the Intention-to-Treat (ITT) to distinguish from the actual treatment status after some farmers switched MCPs. Farmer-level variables based on administrative data include measures of milk quality[62] and milk prices.

Additional individual farmer-level covariates are from a survey conducted in October-November 2015. We surveyed all farmers delivering to thirteen selected MCPs – six (out of seven) of the MCPs that were upgraded during the study period, and seven non-upgraded MCPs. Out of a sample of 1,351 farmers, no less than 1,335 farmers participated in our survey (98.8% compliance rate). The survey includes questions on demographics, household assets,

[61] Payment groups are groups of, on average, six producers (with a maximum of nineteen producers) who pool their individual milk production and are paid based on their average quality. It is evident that payments based on average quality attenuates incentives to invest in quality, as the costs of quality-enhancing initiatives are private and the benefits are shared with other members of the payment group.

[62] These include the Total Solids (TS) content and the Freezing Point (FP) as measures for the compositional quality of milk, and the Total Plate Count (TPC) as measures for hygienic quality. For more details on quality measures we refer to Chapter 4.

number of dairy cows, farm assets, labor and trust.[63] We also know the member ID of farmers, so we can link survey data to cooperative administrative data discussed above.[64]

Table 6.1 summarizes our outcome variables, which are all derived from follow-up cooperative administrative data. First, we assess differences in milk prices received in the months after the opening of the last upgraded MCP in January 2018. Second, we explore differences in quantities delivered. Since the start of the intervention, the cooperative transitioned from measuring milk quantity in liters to kilograms at both upgraded and non-upgraded MCPs. As the latter is more accurate, we only use measures after all MCPs transitioned. Hence we measure quantity delivered by the sum of quantities delivered during the last two months of 2018. Third, we explore whether the intervention affected ‘attrition’, or the probability that smallholders exited the industry. This is measured by a dummy indicating whether or not the farmer delivered milk to the cooperative in 2018.

Table 6.1. Overview of Outcome Variables

Outcome variable	Unit	Type	Period	Sample[1]
Price	IDR[2]	Mean	Feb-Dec 2018	Delivered milk during Feb-Dec-2018
Quantity	kg	Total	Nov-Dec 2018	Delivered milk during Feb-Dec-2018
Delivered milk	-	Dummy	Feb-Dec 2018	All

[1] Of farmers surveyed at baseline
[2] IDR = Indonesian Rupiah

The MCP-level inclusiveness analysis includes all MCPs of the cooperative, and variables are constructed based on records of all farmers delivering milk in 2014. The farmer sample consists of all farmers delivering milk in 2014 and participating in our baseline survey.[65] For the analysis of prices received and quantities delivered at endline we restrict the analysis to the farmers that delivered at least some milk during the endline period.

[63] For details, please refer to Appendix Table A6.1

[64] When farmers indicated that they had more than one member ID, we track only the first member ID over time, as extra member IDs are more likely to be held temporarily. In December 2017 and January 2018, slightly more than two years after the baseline survey, less than 1% of the main IDs was held by a different household, either temporarily or permanently.

[65] We exclude from the study of endline differences one board member and sixteen farmers who participated in another training program (but including these observations does not affect any of the outcomes).

Table 6.2 provides a timeline of the intervention and data collection efforts. Observe that the baseline was conducted *after* the opening of the first upgraded MCP. Further, the endline period starts very soon after the opening of the final two upgraded MCPs, leaving little time for impact. These issues will be addressed in robustness checks.

Table 6.2. Timing of Intervention and Data Collection

	2014	2015	2016	2017	2018
MCP upgrades:					
- 1st					
- 2nd					
- 3rd					
- 4th					
- 5th					
- 6th					
Data collection:					
- Baseline administrative					
- Baseline survey					
- Endline administrative					

Intervention training periods are indicated in gray

6.4. Who is reached by the intervention?

To analyze inclusiveness we first compare baseline MCP-level characteristics between upgraded and non-upgraded MCPs. Next, we compare baseline farmer-level characteristics for members of upgraded and non-upgraded MCPs. This includes data on demographics, household assets, number of dairy cows, farm assets, labor, trust and milk quality, quantity, prices and income. Given the small number of MCPs in our data set, we correct for clustering at MCP level using a wild bootstrap-t procedure as proposed by Cameron, Gelbach, and Miller (2008) throughout this paper.

Table 6.3 shows the results of the *MCP-level* balance analysis. Prior to the intervention, selected MCPs were significantly larger than non-selected MCPs in terms of the number of farmers delivering and the total quantity of milk delivered. It is not surprising that these two variables are closely correlated ($r = 0.949$). Ranking MCPs by the number of farmers delivering, the four largest MCPs were upgraded, together with the 10th, 13th and 19th largest (out of a total of 31 MCPs).

Table 6.3. Inclusiveness of the Intervention (MCP Level Variables)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Number of farmers	7	125.1	50.9	24	61.4	34.8	0.018
Total milk quantity (1,000,000 L)	7	1.569	0.713	24	0.713	0.362	0.008
Milk quantity per farmer (1,000 L)	7	8.812	2.169	24	6.885	1.160	0.032
Milk compositional quality TS (%)	7	11.71	0.20	24	11.69	0.14	0.842
Milk compositional quality FP (>-0.520°C)	7	0.219	0.103	24	0.259	0.112	0.396
Milk compositional quality FP (>-0.500°C)	7	0.024	0.023	24	0.024	0.025	0.990
Milk hygienic quality TPC (bonus)	7	0.506	0.155	24	0.486	0.136	0.744
Milk price (1,000 IDR)	7	4.104	0.092	24	4.085	0.076	0.600
Milk income (1,000,000 IDR)	7	33.505	8.563	24	26.022	4.572	0.036
Board member	7	0.143	0.378	24	0.042	0.204	0.766

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

We do not find significant evidence that the selection of MCPs for upgrading is driven by cooperative board membership, so there is no evidence of local elite capture. However, based on this result, we cannot rule out that an exception is the 13th largest MCP, where two cooperative board members deliver their milk on a daily basis.

From a poverty (or distributional) perspective, more important than the MCP-level analysis is a comparison at the *farmer-level*.[66] Table 6.4 provides the farmer-level balance results. Farmers at upgraded MCPs were significantly more wealthy and owned more cows than farmers at non-upgraded MCPs. They also score higher on related variables, such as the Progress out of Poverty Index, ownership of household and farm assets, production of milk quantity, and milk income.

6

[66] The data in Appendix Table A6.2 indicate that the 13 MCPs included in our farmer-level analysis (upgraded and non-upgraded) are fairly representative of the universe of 31 MCPs supplying to the cooperative. When we re-estimate Table 6.3 for MCPs selected for the baseline survey, we find that MCP level variables differ substantially only for the number of farmers and the total milk quantity delivered (and not for milk quantity per farmer and total milk income). For these variables, differences are slightly smaller within the subset of MCPs selected for surveys than within those that were not selected.

Table 6.4. Inclusiveness of the Intervention (Farmer Level Variables)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	685	44.05	12.34	605	45.27	12.23	0.182
Gender of respondent	685	0.164	0.370	605	0.099	0.299	0.082
Junior high school or higher	685	0.327	0.469	605	0.276	0.447	0.572
Household asset index	685	4.258	1.759	605	3.914	1.547	0.040
International Wealth Index	685	67.85	12.98	605	61.65	12.70	0.012
Progress out of Poverty Index	685	39.26	8.63	605	37.22	8.75	0.056
Number of dairy cattle total	685	5.505	4.151	605	3.909	2.882	0.010
Farm asset index	685	2.790	0.781	605	2.233	0.806	0.060
Number of non-family fulltime workers	685	0.091	0.387	605	0.051	0.329	0.212
Number of non-family parttime workers	685	0.023	0.201	605	0.007	0.081	0.332
Number of family fulltime workers	685	1.223	0.756	605	1.185	0.655	0.488
Number of family parttime workers	685	0.619	0.554	605	0.593	0.534	0.628
Milk quantity (1,000 L)	685	10.237	8.727	605	8.138	5.982	0.102
Milk quantity including extra IDs (1,000 L)	685	10.393	8.878	605	8.592	7.852	0.214
Milk compositional quality TS (%)	685	11.72	0.29	605	11.68	0.27	0.634
Milk compositional quality FP (>-0.520°C)	685	0.209	0.236	605	0.237	0.249	0.642
Milk compositional quality FP (>-0.500°C)	685	0.018	0.058	605	0.011	0.042	0.350
Milk hygienic quality TPC (bonus)	685	0.514	0.292	605	0.526	0.311	0.800
Milk price (1,000 IDR)	685	4.131	0.154	605	4.114	0.153	0.688
Milk income (1,000,000 IDR)	685	38.92	33.42	605	30.83	23.08	0.114
Milk income including extra IDs (1,000,000 IDR)	685	39.52	34.01	605	32.58	30.35	0.218
Payment group size	685	6.040	3.890	605	6.925	3.867	0.648
Distance to MCP (m)	685	849.9	806.2	605	757.7	930.1	0.656
Trust index[1]	665	3.610	0.368	599	3.666	0.367	0.214
Children want to take over	685	0.707	0.362	605	0.681	0.337	0.468

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

6

Figure 6.1 shows how the distribution of cows differs across members of upgraded and non-upgraded MCPs. The distribution of farmers at upgraded MCPs lies more ‘to the right’ of the distribution of farmers at non-upgraded MCPs, again showing that smallholders with more cows are more likely to receive treatment. The graph further shows that the probability of being reached linearly increases with the number of cows over the range of common values, and more than doubles over this range.

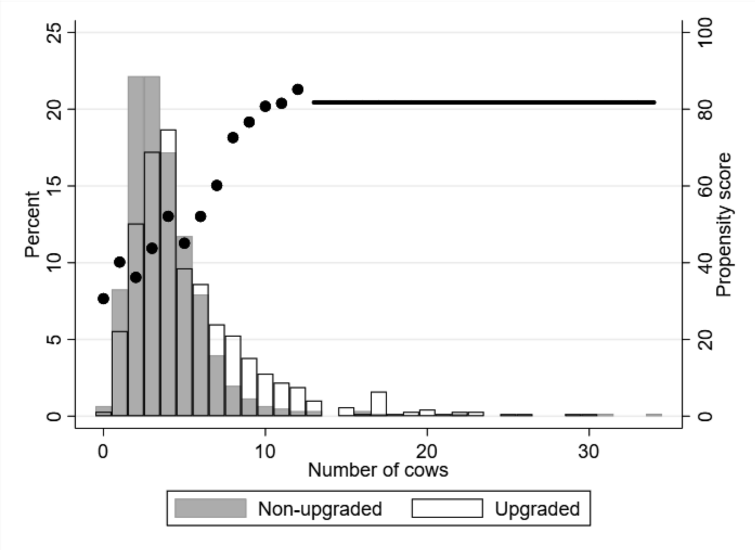

Figure 6.1. Histogram of number of cows by intervention status (ITT) overlaid by scatterplot of propensity score. Because of the low number of farmers with many cows, farmers with 13 or more cows were aggregated for the calculation of the propensity score.

We have performed some additional analyses to probe the robustness of these results. Since baseline surveys were conducted after the opening of the first upgraded MCP, one might fear that imbalance reflects impacts of the intervention rather than pre-existing differences. We therefore re-estimated the results in Table 6.4 and Figure 6.1 while excluding the first upgraded MCP. Results are included in Appendix Table A6.3 and Figure A6.1, and are similar to the results above (but *p*-values are slightly higher in this reduced sample).

6.5. Winners and losers of the PPP intervention?

The multinational dairy processor pays the cooperative based on the quantity and quality of the milk that is supplied. The firm rewards quality and applies an elaborate system with a base price per unit of milk, combined with bonuses for compositional and hygienic quality. The cooperative applies a slightly augmented version of this price scheme when purchasing milk from groups or individual suppliers, but the ‘pass-through’ of bonuses is nearly one-on-one. In other words, the price matrix of the cooperative follows the price matrix of the firm, presumably in an effort to incentivize farmers to produce high-quality milk as desired by the firm.

Importantly, the cooperative offers the same price scheme to all members, across all MCPs. Distributional issues may emerge if the multinational firm alters its price scheme – exactly as it did following the upgrading of selected MCPs. As the firm seeks to purchase high-quality milk, it incentivizes the production of milk with a high solid content and low bacteria contamination. Chapter 5 establishes that, as quality increased, *the firm paid a higher bonus for quality, but lowered the base price*.

Specifically, Chapter 5 studied prices paid by the multinational firm over time, and found that increased hygienic (and compositional) quality resulted into higher prices *in the short-run only*. Within a period of months, however, these price premiums disappeared as the buyer did not maintain a competitive base price, but instead decreased its base price relative to another processor. As a result, prices converged to a level where the cooperative was (nearly) indifferent between selling to the PPP firm and this other processor. Consequently, smallholder producers gain very little on average – essentially they are kept at their reservation value, as determined by the price of milk offered by a rival processing firm. But not everybody is affected the same way.

To evaluate the distributional effects of the intervention we need to compare similar farmers from upgraded and non-upgraded MCPs. To create such comparison groups, we employed a Coarsened Exact Matching procedure as discussed by Iacus, King, and Porro (2012). This procedure re-weights observations from non-upgraded MCPs to mimic the distribution of observations at upgraded MCPs for pre-selected matching variables. We exactly matched farmers on gender, presence in one of six equidistant intervals of the International Wealth Index, number of cows, and farm assets. For each outcome regression, we first employ our matching procedure on the relevant sample (as listed in Table 6.1). Appendix Tables A6.4 and A6.5 show that matching resulted in very similar comparison groups for all samples.[67] To gauge the distributional effect of the intervention, we regress outcome variables on the intervention indicator, including co-variates listed in Table 6.4 as baseline controls in some of the models.

Table 6.5 presents the resulting endline differences for the outcome variables listed in Table 6.1, as well as summary statistics for these outcomes for farmers at non-upgraded MCPs.

[67] We only find that the difference on the number of non-family full-time workers was significantly larger for farmers from upgraded MCPs, but the absolute difference was still small.

Columns (1-2) show that the price received by farmers from upgraded MCPs is about 5% higher than prices received by other farmers. This reflects pass-through of the incentives provided by the multinational firm. Farmers at upgraded MCPs are in a better position to respond – given the training they received and the absence of incentive-dilution due to group-level quality measurement. Given the costs of inputs, and the resulting small margins, these price differences are substantial. In contrast, we find very small and statistically insignificant differences in terms of the quantity of milk delivered, or on the probability of delivering at least some milk during the endline period (but statistical power is limited).

Table 6.5. Endline Differences

	Outcome variables					
	(1)	(2)	(3)	(4)	(5)	(6)
	Price (IDR)	Price (IDR)	Quantity (kg)	Quantity (kg)	Delivered milk	Delivered milk
Upgraded (ITT)	212.7***	226.5***	-42.70	0.795	0.0164	0.0131
- Clustered SE	(32.1)	(27.4)	(167.3)	(84.9)	(0.0426)	(0.0338)
- Bootstrapped p-value	0.000	0.000	0.832	0.954	0.738	0.722
Baseline controls	No	Yes	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer	Farmer	Farmer
Sample	Active	Active	Active	Active	All	All
MCPs	13	13	13	13	13	13
Observations	755	755	755	755	962	962
Mean of non-upgraded	4703.5	4703.5	1677.2	1677.2	0.8089	0.8089
SD of non-upgraded	162.6	162.6	1199.6	1199.6	0.3936	0.3936

Standard errors clustered at MCP level in parentheses

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

* p<0.10, ** p<0.05, *** p<0.01

As robustness checks we again exclude the first upgraded MCP, and exclude the final two upgraded MCPs. Results are reported in Tables A6.6 and A6.7. Excluding these MCPs from the dataset does not materially affect results for the price variable. Point estimates on quantities delivered and attrition are larger, but remain insignificant at conventional levels. The directions of the effects, however, suggest that farmers at upgraded MCPs have increased their milk production and are more likely to have continued as a dairy farmer relative to farmers at non-upgraded MCPs.

What has been the effect of the PPP intervention? Consider the following three pieces of evidence. First, the PPP intervention did not affect *average* prices paid to smallholders, as the downward adjustment in the base price more or less offsets the increase in bonuses (Chapter 5). Second, the PPP intervention reaches MCPs serving relatively wealthy (and productive) farmers. Third, the PPP intervention raised prices for farmers supplying to upgraded MCPs, relative to prices for farmers supplying to non-upgraded MCPs. Taking this evidence together suggests that the PPP intervention not only increased inequality among our sample of smallholder producers, it also impoverished the poorest subsample of smallholders relative to the counterfactual without the intervention.

As a limitation to this study, we have data for just one cooperative, so we are not able to detect potential spillover effects in this study. If (i) the intervention at upgraded MCPs indirectly helped farmers at non-upgraded MCPs to increase milk quality, and (ii) increased milk quality is also rewarded by other processors, then farmers at non-upgraded MCPs might still have benefitted indirectly from the intervention.

6.6. Discussion and conclusion

The Aid for Trade intervention that we study reaches relatively large MCPs and wealthy farmers. This reflects a greater return to investment for the firm: the upgraded MCPs reach many farmers, and the investments in training affect the production of many cows. The cooperative context in which the MCP-level intervention was rolled out necessarily implies that some relatively poor farmers (in terms of wealth and the number of cows owned) were also reached. However, our data suggest that the intervention has increased economic inequality within the cooperative membership, and transferred income from relatively poor member groups to more wealthy ones as a result of the intervention. This is where the efficiency-equity (or efficiency-poverty) trade-off is particularly strong.

The existence of such trade-offs is well-known. In the context of subsidized microfinance the debate is about 'outreach' (to poor borrowers) versus 'financial sustainability' (for the lender) (D'Espallier et al. 2017, Mia and Lee 2017). In the context of health interventions involving subsidized health inputs, the debate is about wasteful 'over-inclusion' (due to subsidies) versus 'over-exclusion' (rationing the poor out of these markets – see Dupas 2014). It should be no surprise that similar dilemmas emerge in the context of subsidized interventions in agro-food value chains. PPP interventions, supported with public funding, where private firms are behind the driving wheel, are not necessarily beneficial for those most in need of assistance.

Appendix

Table A6.1. Details on Survey Variables

Household asset index	Sum of dummies indicating whether the houshold owns at least one:
	- Watch
	- Mobile telephone
	- Smart-phone
	- Bank account
	- Radio
	- Television
	- Refridgerator
	- Freezer
	- Computer
	- Bicycle
	- Motor
	- Car of truck
	- Generator
	- Solar panel
	- Gas cilinder
International Wealth Index	See Smits and Steendijk (2015)
Progress out of Poverty Index	See Schreiner (2012)
Farm asset index	Sum of dummies indicating whether the houshold owns at least one:
	- Barn or cowshed
	- Chopper for cutting the grass
	- Animal-drawn cart
	- Milk can
	- Milking machine
	- Irrigation equipment
Trust index	Mean of 5-point Likert scores on trust in:
	- The local dairy cooperative
	- The local government
	- The processor that was part of the PPP
	- The other processor that buys a lot of milk from the cooperative
	- Other dairy farmers in your payment group
	- Other dairy farmers in your farm group
	- Other farmers in general
Children want to take over	Do you think your children would want to join or take over your dairy farming business at some stage?
	- 0 if No
	- 0.5 if Uncertain
	- 1 if Yes

Table A6.2. Inclusiveness of the Intervention (MCP Level Variables; Excluding MCPs Not Surveyed)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Number of farmers	6	117.5	51.2	7	72.4	37.9	0.117
Total milk quantity (1,000,000 L)	6	1.479	0.736	7	0.860	0.427	0.085
Milk quantity per farmer (1,000 L)	6	8.739	2.367	7	6.904	1.006	0.099
Milk compositional quality TS (%)	6	11.69	0.21	7	11.70	0.15	0.900
Milk compositional quality FP (>-0.520°C)	6	0.239	0.097	7	0.254	0.128	0.833
Milk compositional quality FP (>-0.500°C)	6	0.027	0.023	7	0.016	0.020	0.361
Milk hygienic quality TPC (bonus)	6	0.469	0.134	7	0.515	0.157	0.569
Milk price (1,000 IDR)	6	4.091	0.093	7	4.095	0.074	0.909
Milk income (1,000,000 IDR)	6	33.128	9.316	7	26.193	4.017	0.113
Board member	6	0.167	0.408	7	0.000	0.000	0.439

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

6

Table A6.3. Inclusiveness of the Intervention (Farmer Level Variables; Excluding 1st Upgraded MCP)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	496	44.12	11.80	605	45.27	12.23	0.304
Gender of respondent	496	0.171	0.377	605	0.099	0.299	0.134
Junior high school or higher	496	0.270	0.444	605	0.276	0.447	0.950
Household asset index	496	4.137	1.680	605	3.914	1.547	0.104
International Wealth Index	496	67.06	12.72	605	61.65	12.70	0.008
Progress out of Poverty Index	496	38.16	8.25	605	37.22	8.75	0.082
Number of dairy cattle total	496	4.990	3.731	605	3.909	2.882	0.038
Farm asset index	496	2.677	0.714	605	2.233	0.806	0.108
Number of non-family fulltime workers	496	0.063	0.315	605	0.051	0.329	0.466
Number of non-family parttime workers	496	0.018	0.184	605	0.007	0.081	0.746
Number of family fulltime workers	496	1.252	0.773	605	1.185	0.655	0.268
Number of family parttime workers	496	0.639	0.551	605	0.593	0.534	0.402
Milk quantity (1,000 L)	496	8.997	7.126	605	8.138	5.982	0.214
Milk quantity including extra IDs (1,000 L)	496	9.190	7.347	605	8.592	7.852	0.420
Milk compositional quality TS (%)	496	11.73	0.31	605	11.68	0.27	0.738
Milk compositional quality FP (>-0.520°C)	496	0.243	0.256	605	0.237	0.249	0.950
Milk compositional quality FP (>-0.500°C)	496	0.023	0.067	605	0.011	0.042	0.190
Milk hygienic quality TPC (bonus)	496	0.473	0.293	605	0.526	0.311	0.464
Milk price (1,000 IDR)	496	4.119	0.163	605	4.114	0.153	0.928
Milk income (1,000,000 IDR)	496	34.15	27.45	605	30.83	23.08	0.228
Milk income including extra IDs (1,000,000 IDR)	496	34.89	28.30	605	32.58	30.35	0.432
Payment group size	496	6.908	4.109	605	6.925	3.867	0.960
Distance to MCP (m)	496	970.8	846.0	605	757.7	930.1	0.416
Trust index[1]	483	3.609	0.380	599	3.666	0.367	0.282
Children want to take over	496	0.709	0.360	605	0.681	0.337	0.516

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

6

Table A6.4. Balance at Baseline in Matched Sample (Active at Endline)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	381	44.59	12.41	374	44.36	10.70	0.792
Gender of respondent	381	0.076	0.266	374	0.076	0.266	1.000
Junior high school or higher	381	0.278	0.449	374	0.318	0.467	0.648
Household asset index	381	3.971	1.494	374	4.272	1.589	0.182
International Wealth Index	381	66.19	10.98	374	65.81	10.71	0.802
Progress out of Poverty Index	381	38.45	8.26	374	38.58	8.68	0.948
Number of dairy cattle total	381	4.349	2.122	374	4.349	2.122	0.993
Farm asset index	381	2.701	0.640	374	2.701	0.641	1.000
Number of non-family fulltime workers	381	0.058	0.284	374	0.022	0.168	0.048
Number of non-family parttime workers	381	0.018	0.169	374	0.006	0.076	0.268
Number of family fulltime workers	381	1.131	0.672	374	1.201	0.576	0.310
Number of family parttime workers	381	0.617	0.533	374	0.587	0.511	0.694
Milk quantity (1,000 L)	381	8.614	5.538	374	8.703	5.169	0.964
Milk quantity including extra IDs (1,000 L)	381	8.781	5.812	374	8.960	5.681	0.876
Milk compositional quality TS (%)	381	11.73	0.29	374	11.73	0.27	1.000
Milk compositional quality FP (>-0.520°C)	381	0.209	0.236	374	0.209	0.242	0.958
Milk compositional quality FP (>-0.500°C)	381	0.017	0.051	374	0.008	0.032	0.256
Milk hygienic quality TPC (bonus)	381	0.517	0.290	374	0.576	0.308	0.456
Milk price (1,000 IDR)	381	4.131	0.156	374	4.139	0.149	0.870
Milk income (1,000,000 IDR)	381	32.73	21.23	374	33.15	19.90	0.956
Milk income including extra IDs (1,000,000 IDR)	381	33.37	22.29	374	34.14	21.96	0.868
Payment group size	381	6.223	3.992	374	6.698	3.527	0.740
Distance to MCP (m)	381	887.5	823.6	374	786.2	806.7	0.714
Trust index[1]	370	3.610	0.374	371	3.646	0.352	0.366
Children want to take over	381	0.690	0.355	374	0.656	0.355	0.478

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

6

Table A6.5. Balance at Baseline in Matched Sample (All)

	Upgraded (ITT)			Non-upgraded (ITT)			Diff
	N	mean	sd	N	mean	sd	p
Age of respondent	475	44.33	12.48	487	45.31	11.55	0.368
Gender of respondent	475	0.082	0.275	487	0.082	0.275	1.000
Junior high school or higher	475	0.282	0.450	487	0.311	0.463	0.732
Household asset index	475	3.931	1.493	487	4.202	1.566	0.168
International Wealth Index	475	65.79	11.17	487	65.54	10.74	0.818
Progress out of Poverty Index	475	38.14	8.11	487	38.48	8.86	0.762
Number of dairy cattle total	475	4.147	2.132	487	4.147	2.132	1.000
Farm asset index	475	2.674	0.640	487	2.674	0.640	0.961
Number of non-family fulltime workers	475	0.048	0.259	487	0.043	0.264	0.804
Number of non-family parttime workers	475	0.015	0.152	487	0.008	0.088	0.528
Number of family fulltime workers	475	1.160	0.669	487	1.220	0.586	0.360
Number of family parttime workers	475	0.627	0.545	487	0.568	0.511	0.330
Milk quantity (1,000 L)	475	8.326	5.588	487	8.333	4.976	0.932
Milk quantity including extra IDs (1,000 L)	475	8.479	5.811	487	8.570	5.386	0.864
Milk compositional quality TS (%)	475	11.71	0.30	487	11.72	0.27	0.972
Milk compositional quality FP (>-0.520°C)	475	0.222	0.244	487	0.223	0.248	1.000
Milk compositional quality FP (>-0.500°C)	475	0.020	0.065	487	0.009	0.037	0.212
Milk hygienic quality TPC (bonus)	475	0.501	0.291	487	0.553	0.306	0.460
Milk price (1,000 IDR)	475	4.122	0.159	487	4.129	0.149	0.870
Milk income (1,000,000 IDR)	475	31.58	21.47	487	31.64	19.09	0.918
Milk income including extra IDs (1,000,000 IDR)	475	32.16	22.32	487	32.55	20.74	0.852
Payment group size	475	6.295	3.971	487	6.612	3.514	0.866
Distance to MCP (m)	475	908.2	824.2	487	824.9	885.8	0.674
Trust index[1]	463	3.616	0.371	483	3.644	0.358	0.596
Children want to take over	475	0.695	0.358	487	0.669	0.349	0.528

Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)

[1] For later regression analysis, missings are set to the mean of farmers with the same upgrade status (ITT)

6

Table A6.6. Endline Differences (1st Upgraded MCP Excluded)

	Outcome variables					
	(1)	(2)	(3)	(4)	(5)	(6)
	Price (IDR)	Price (IDR)	Quantity (kg)	Quantity (kg)	Delivered milk	Delivered milk
Upgraded (ITT)	206.5***	220.7***	-96.85	2.219	0.00574	0.00975
- Clustered SE	(32.1)	(26.6)	(165.5)	(101.7)	(0.0446)	(0.0415)
- Bootstrapped p-value	0.002	0.002	0.582	1.000	0.924	0.812
Baseline controls	No	Yes	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer	Farmer	Farmer
Sample	Active	Active	Active	Active	All	All
MCPs	12	12	12	12	12	12
Observations	653	653	653	653	844	844
Mean of non-upgraded	4705.2	4705.2	1620.0	1620.0	0.8013	0.8013
SD of non-upgraded	161.8	161.8	1170.2	1170.2	0.3994	0.3994

Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

Table A6.7. Endline Differences (MCPs Upgraded in 2018 Excluded)

	Outcome variables					
	(1)	(2)	(3)	(4)	(5)	(6)
	Price (IDR)	Price (IDR)	Quantity (kg)	Quantity (kg)	Delivered milk	Delivered milk
Upgraded (ITT)	208.9***	226.3***	47.50	85.56	0.0378	0.0369
- Clustered SE	(33.8)	(29.6)	(161.7)	(66.8)	(0.0368)	(0.0282)
- Bootstrapped p-value	0.000	0.000	0.774	0.228	0.317	0.201
Baseline controls	No	Yes	No	Yes	No	Yes
Unit of analysis	Farmer	Farmer	Farmer	Farmer	Farmer	Farmer
Sample	Active	Active	Active	Active	All	All
MCPs	11	11	11	11	11	11
Observations	659	659	659	659	840	840
Mean of non-upgraded	4705.7	4705.7	1678.1	1678.1	0.8155	0.8155
SD of non-upgraded	166.8	166.8	1198.6	1198.6	0.3883	0.3883

Standard errors clustered at MCP level in parentheses
Bootstrapped p-values obtained with wild bootstrap using Rademacher weights and imposing the null hypothesis as proposed by Cameron, Gelbach and Miller (2008)
* p<0.10, ** p<0.05, *** p<0.01

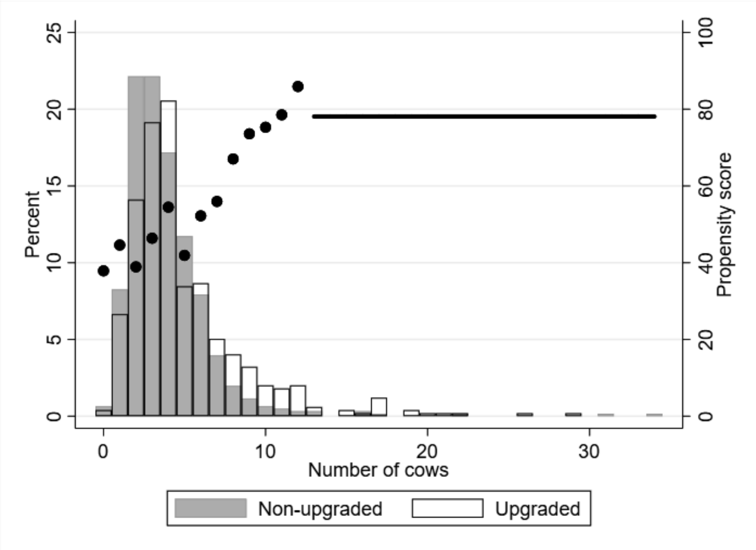

Figure A6.1. Histogram of number of cows by intervention status (ITT) overlaid by scatterplot of propensity score (excluding 1st upgraded MCP). Because of the low number of farmers with many cows, farmers with 13 or more cows were aggregated for the calculation of the propensity score.

6

7. General Discussion

Mark Treurniet

7

7.1. Introduction

Development agencies frequently involve the private sector through public-private partnerships (PPPs) to leverage the private sector's expertise, efficiency and capital, and together realize development goals. However, little is known about the contribution that PPPs are able to make to development goals. The General Introduction of this thesis therefore defined the main research question of this thesis as follows:

"What is the potential for public-private partnerships to support food safety and farmer incomes in food value chains in lower middle-income countries?"

The General Introduction also presented a summary of the evidence presented in the following five Chapters. This General Discussion will synthesize the available evidence, provide recommendations for public-private interventions in food value chains, and discuss considerations for further research.

7.2. Synthesis

Taken together, the existing evidence and the additional evidence presented in this thesis suggest that the alignment of private interests with development goals is key to realize intended development outcomes. When private interests are well-aligned with development goals, PPPs may be an effective organizational tool to spur development. Instead, when private interests and development goals are misaligned, intended outcomes are unlikely to be reached (Hart and Holmström 1987, Poulton and Macartney 2012).

7.2.1. Potential for PPPs: Alignment of private interests with development goals

I first discuss some areas where private interests seem to be well-aligned with development goals, and where PPPs may be an effective tool to spur development. The potential especially exists in the realization of food safety (SDG 2).

Stimulating demand for safe food

For a market-driven solution to food safety hazards, consumers need to value food safety or at least attributes that correlate with food safety, like perhaps certain dimensions of food quality or the ability to process it into higher value products. It is often found that consumers do look for alternatives when they are aware of food safety risks (Ahmed et al. 2006). Meanwhile, many studies obtain high estimates for the Willingness To Pay for safe food, but elicit preferences right after food safety issues were discussed with participants (Hoffmann, Moser, and Saak 2019), and Chapter 2 suggests that this can lead to upward-biased estimates.

In low income areas with moderate food safety risks, it might actually be very challenging to sustainably increase consumers' Willingness To Pay for safe food. For example, Hoffmann, Moser, and Herrman (2019) find that a food safety campaign affected the sales of a safe maize flower brand, but this effect disappeared after the campaign stopped. When an additional discount was offered, the impact lasted longer, but eventually also faded. The authors suggest that consumers may assume that food is generally safe, and that positive signals about safety therefore do not change much. This basic assumption is of course likely to change under extreme circumstances.

While increasing demand for safe food may be challenging, the interests of private companies can be aligned with the objective to increase demand for food safety if these companies are able to produce safe alternatives in a cost-effective way. Depending on the institutional context, PPPs can therefore be an effective tool to address the demand for food safety.

Incentivizing the supply of safe food

In contexts where markets are willing to reward the safety and quality[68] of produce, introduction of price incentives offered on the basis of third-party measurement has been found to increase the safety and quality of food delivered to the market (Saenger et al. 2013, Bernard et al. 2017). Among subsistence farmers, the introduction of market premiums increases the adoption of food safety technology at the extensive margin (whether or not one adopts) in cases where adoption was a binary decision (Hoffmann, Magnan, et al. 2018), or an unrealistically large premium was offered (Hoffmann and Jones 2018).

This thesis provides additional insights on two points. First, perhaps counterintuitively, Chapter 3 showed that modest price premiums that are too low to increase adoption at the extensive margin, can still increase adoption at the intensive margin (the intensity of adoption) among subsistence farmers. Farmers may value safe food for home consumption, and therefore adopt some food safety technology. In bad years, all safe food will be consumed at home, but in good years the excess safe food will be delivered to the market. With the introduction of a market

[68] See the General Introduction for a discussion on the concepts of food safety and food quality.

premium, the expected commercialization benefit increases, and therefore households adopt more. Since in bad years farmers may still prefer to consume all safe food at home, small market incentives have the potential to contribute not only to increased availability of safe food in the market, but also to the safety of home consumption.

To better understand how modest market premiums interact with the household's value premium for safe home consumption, it might be useful to recall from the model of Chapter 3 that the introduction of a market premium only increased the adoption of food safety technology at the intensive margin if the value premium for safe home consumption is neither too small (since the modest premium is insufficient to fully compensate for the costs of adoption) nor too high (otherwise the household already adopts maximally even without a market premium). This can explain not only why Chapter 3 found the market premium to increase adoption only at the intensive margin, but also why Chapter 2 did not find the market incentive to increase adoption in the surveyed sample. Modest market premiums for safe food thus seem to have the largest effect on food safety technology for intermediate levels of the household's value premium for safe home consumption.

Second, in all the studies cited above, quality is measured by independent parties. If measurements from buyers are less trusted, impacts on food quality might be lower (Saenger, Torero, and Qaim 2014, Abate and Bernard 2017). Chapter 4 showed that even if samples are tested by the buyer, quality incentives have the potential to increase quality delivered by smallholder farmers, at least in the context of dairy, where farmers receive signals on quality measurements on a regular basis.

Even if households value safe food, markets for safe food might not exist if certain characteristics are unobservable and it is hard for sellers to build up a credible reputation (Chapter 3). In such cases, PPPs can introduce independent food safety certification, thus helping the safety of their product to be signaled to consumers, so that premiums can be paid for safe food and markets for safe food arise. If market premiums are sufficiently high, certification schemes can ultimately be financed by the market, but initial public investments might be needed to get the system off the ground.

The rise of markets for safe foods may have the unintended consequence that the safety of food in regular markets decreases. To avoid this, it is critical that incentives are also passed on to aggregators, traders and farmers, so that they are incentivized to increase the quality of production. If market premiums exist downstream in the value chain, PPPs could help to pass on incentives to more upstream actors by introducing measurement and certification systems at intermediate markets. If market premiums are sufficiently high, the running costs of such systems can be financed by the market, but public investments could again be needed for the introduction of these systems.

Since increasing quality can add value to the product, private sector interests are generally likely to be aligned with the development goal to increase the safety of food. However, if some actors become worse off in their own experience, measurement and certification systems may revoke resistance and eventually collapse. For example, this might happen if incentives are framed and experienced as penalties (Chapter 4), or if increased transparency decreases profit margins of intermediaries and PPPs cannot make these intermediaries redundant (Bernard et al. 2017). The latter example concerns imperfect competition, and I will discuss later how other interventions might handle problems of imperfect competition. Finally, consumers' trust in institutions is context-specific, and can affect demand for certified products, so it will depend on the context whether public or private will be more effective (Hoffmann, Moser, and Saak 2019).

Linking high-potential smallholder farmers to modern value chains I

PPPs might find ways to help smallholder farmers to satisfy and keep up with increasing quality requirements (Chapter 4), and thereby enable them to deliver to modern value chains. PPPs might especially be able and willing to help farmers and groups that produce relatively large amounts of food, as it likely is most profitable to train these farmers (Chapter 6). Of course, these farmers and groups are the most important to reach when one wants to improve food safety.

7.2.2. Limitations to PPPs: Misalignment of private interests with development goals

Private interests might be less aligned with other development goals, and especially poverty reduction (SDG 1).

Agricultural extension

Private value chain actors can jointly increase profits and contribute to development goals by providing agricultural extension, and thus helping farmers to adopt specific technologies sold by this actor, or produce better output to be bought by this actor. As private actors (i) may have important knowledge, (ii) may be able to deliver agricultural extension at low cost, and (iii) may be willing to co-finance agricultural extension, donors intend to contribute to development goals by providing public funding to private companies to scale-up agricultural extension (see Chapters 4-6 for an example). However, private firms may prioritize their private interests over development goals. For example, a report by De Brauw et al. (2018) documents the impacts from a training delivered by a NGO and promotions delivered by a private sector input provider, and shows that the first increased general knowledge on inputs and practices, while second specifically increased knowledge on the use of fertilizers, which could be bought from the input provider. While the private promotions increased fertilizer adoption and thus contributed to the input provider's revenues, it has remained an open question which intervention was more effective to boost productivity. As another example, the dairy processor from Chapters 4-6 valued hygienic quality more than competing dairy processors, and this might have been the reason that they led the PPP to support smallholder farmers increase quality in this specific dimension, even though farmers could perhaps have benefitted more from other interventions. Whether private value chain actors are the most appropriate candidate to deliver publicly funded agricultural extension again depends on the alignment of private sector interests and development goals.

Increasing farmer incomes

Existing evidence suggests that agricultural crop markets are rather competitive, at least at the level of traders (Dillon and Dambro 2017). On the contrary, Chapter 5 shows that processors and exporters may exhibit significant market power over quality rents, and may be able to extract intervention rents. Even though buyers may include premiums for high-quality food in their price matrices, farmers will not benefit from higher prices for higher quality food if buyers have full market power over quality rents.

Further research is needed to learn how common market power and rent-extraction are, and how PPPs change underlying market structures. In the next section I will discuss how interventions in value chains could potentially address market power issues. In the subsequent section I will discuss how further research could help to further improve our understanding on this issue.

Linking low-potential smallholder farmers to modern value chains II

Most PPPs are generally not active in the least developed and fragile countries (Bouma and Berkhout 2015). And where PPPs are active, it can be relative costly to help farmers and groups that produce small amounts of food, so that the poorest farmers fall by the wayside when PPPs roll out their interventions (Chapter 6). From an anti-poverty perspective, it is more important to provide support in the poorest places, and to the poorest people. As it is harder to earn a private return to investment in such contexts, PPPs generally appear not to be a good tool to support the poorest people.

7.3. Recommendations for public-private investments in food value chains

The evidence presented in this thesis results into practical recommendations that speak to (i) the type of activities funded, (ii) the selection of implementing partners, (iii) the incentive structure offered by donors to private investors, and (iv) how to create and exploit opportunities for learning.

7.3.1. The type of activities funded

When buyers have substantial market power, producers may not benefit from interventions that help improving the quality of food produced. Instead, Chapter 5 shows that these rents may be extracted by buyers. Especially in developing countries, it might be difficult to directly regulate imperfectly competitive markets. This implies that donors should take market structures into account when designing interventions.

If a buyer has monopsonistic power over rents created in one quality dimension, but not for rents created in another quality dimension, donors can avoid intervention rents to be extracted by creating quality rents in the latter “competitive” dimension. For example, in Chapter 5, the buyer was shown to have market power for the dimension of hygienic quality. While Chapter 5 could not identify whether the buyer also had market power in the dimension of compositional quality, conversations with local agents suggest that it has less market power in this dimension.

Alternatively, interventions could be designed to address the bargaining power of smallholders. The bargaining position of smallholders might be improved in various ways. First, founding a marketing cooperative can increase bargaining power of the producers, and eliminate buyers’ opportunities to differentiate prices paid to different producers. Second, product differentiation or diversification might help to earn higher prices for food that is delivered to new niches or markets, and to improve the outside option and negotiate higher prices for food that continues to be delivered to existing buyers. Third, entry at the buyer level may decrease the bargaining power of the buyer. Fourth, new information technologies may help farmers to get better information on prices, and help them to connect to other sellers.

However, when farmers are faced with various market inefficiencies, buyers may also play a crucial role by providing them with technical assistance, credit and inputs through interlinked contracts. By creating alternative opportunities for the use of these resources and eliminating opportunities for rent-extracting, some of the interventions discussed above may undermine incentives for buyers to offer interlinked contracts (Macchiavello and Morjaria 2017), and cause decreased investments in smallholder farmers and lower farm profitability. As will be discussed later, more research is needed to better understand these dynamics.

7.3.2. The selection of the implementing partners

As we have seen in this thesis, misalignment of private interests and development goals can cause development interventions to fail reaching their intended results. Private actors might design interventions in their private interests, thus not maximizing the contribution to development goals. And given the design of interventions, they might make some parts of the intervention to work better than others. For example, existing buyers may not benefit from quality improvements in certain dimensions or an improved bargaining position of smallholder producers, and might therefore not be motivated to make interventions in this area work. Or buyers may benefit more from training high-potential farmers, and therefore target training away from the poorest farmers. If private interests are misaligned with development goals, it might therefore be better to have independent organizations implementing the publicly funded interventions.

7.3.3. Incentivizing private investors

To realign private sector incentives with development goals and avoid public funds to be spend on projects that do not realize their intended impacts, donors can make their funding contingent on the realization of impacts.

Such a contingency might be contractually arranged through Development Impact Bonds (DIBs) (Center for Global Development and Social Finance 2013). DIBs are a performance-based investment instrument intended to finance development projects. Donors and private investors agree on a shared development goal, how realized impacts will be evaluated, and how payment of funding will depend on the results of this evaluation. Private investors then bear the risk of their investment in interventions and will only earn a return to investment if intended impacts are realized. An example of such a private investor could be a food processor that wants to invest in the capacity of smallholder farmers to produce good quality food.

The use of well-designed DIBs has three main benefits. First, private investors may have better knowledge on the potential of investments in realizing intended results, and will likely not apply for public funding when the expected payoff will be zero. Second, when they do see opportunities for public-private partnership, they are better incentivized to make the investments work for the shared development goal. Third, since measuring impacts is part of the contract, donors will learn about the effectiveness of the chosen strategy.

While DIBs are conceptually easy, several considerations should be taken into account in their design. First, as manipulation-robust and well-identified impact evaluations can be costly, DIBs should be used particularly when enumeration based on inputs delivered is unlikely to lead to the realization of intended impacts. For example, as the dairy processor in Chapters 4-6 seems to benefit from quality improvement, it would be in the interest of this private investor to ensure that inputs will be used effectively to improve quality, even without subsidies that are contingent on impacts. In contrast, Chapter 5 argues that the processor has an incentive to keep prices low, and the project is therefore not guaranteed to contribute to higher farm incomes without the employment of a DIB.

Second, when evaluators aim to estimate the impact on prices, they should take into account the methodological insights from Chapter 5, which are further discussed in the next section. It is especially important that prices will be analyzed at sufficiently aggregate level and access will also be granted to other prices that might proxy for the counterfactual.

Third, when private investors' payoffs depend on results of the impact evaluation, they may want to target investments in such a way that impacts are realized only in the outcome dimension measured for individuals sampled in the period studied, or counterfactual estimates are negatively affected. The design of the evaluation can reduce such manipulation by measuring impact on broad outcomes in whole populations for long periods or keeping any sample selection unknown to the investor, and by ensuring that the counterfactual estimate cannot be negatively affected. The evaluation in Chapter 5 could serve as an example for such a manipulation-robust evaluation: it studies aggregate prices for a relatively long period of time and estimates the counterfactual by a competitor's price, so that it cannot be affected by the private investor.

7.3.4. Creating and exploiting opportunities for learning

When the manipulation-robust and well-identified impact evaluations are too costly, policy makers should strategically create and exploit opportunities to learn about the impact of interventions over time, so that future funding can be redirected to the most effective programs, and new programs can become improved iterations of older ones. However, whether (sufficiently precise) impact estimates can be obtained depends critically on the intervention design. Given the lack of evidence on the effectiveness of public-private interventions and issues raised in this thesis, it would be wise to strategically choose some interventions to be designed in such a way that (sufficiently precise) impact estimates can be obtained, even when this complicates the design and adds transaction costs. As the design of interventions is often within the circle of control of policy makers, I discuss recommendations to increase the evaluability of strategically chosen interventions in this section, while some more technical considerations for further research are discussed in the next section.

First, selection of beneficiaries is often entirely endogenous, so that treated groups are fundamentally different from potential control groups. It is impossible to evaluate the impact of interventions if the parallel trends assumption cannot be made on the basis of empirics or plausibility, or if it is impossible to collect necessary baseline indicators without affecting later impact estimates (Chapter 2). In such cases, the only way to make impact evaluations feasible is to introduce exogenous variation in the intervention status of individuals.

Second, if the parallel trends assumption can be made and baseline data can be collected without affecting later impact estimates, Difference-in-Differences and matching methods might provide reliable impact estimates, like in Chapters 4-6. However, baseline data should then be collected from both the treated and control group before the intervention is announced. This requires timely involvement of evaluators (see Chapters 4 and 6 for a case where this went wrong, but the resulting problems could still be dealt with), and commitment to intervention plans once the sample is fixed, so that the collected data does not become worthless (yes, I also experienced that).

Third, to reduce transaction costs, interventions are often targeted at higher level clusters (e.g. farmer groups, producer organizations, villages or districts), with all individuals within the cluster or an endogenous selection of them being treated. This decreases the statistical power to detect effects between treated and control groups, especially if the number of clusters is small and outcomes are highly correlated within the cluster. This problem and potential solutions will further be discussed in the next section, but the most effective solution to this issue of power is to deliver the intervention at lower level units (e.g. individuals instead of farmer groups, or households instead of villages), provided that spillovers are not too strong between these lower level units. If this is infeasible or undesired, one can consider to add a randomized intervention that encourages some individuals to use the intervention, so that within higher level units, some individuals are more likely to take up the intervention than other individuals. This exogenously induced variation in take-up can then be used to proxy the impact of the intervention (on those that have changed their take-up status based on the encouragement).

7.4. Considerations for further research

The research presented in this thesis has faced various limitations. Many of those limitations have been addressed in the individual Chapters, and in the previous section I already discussed limitations arising from the design of interventions. In this section, I will discuss some general limitations, and subsequently provide considerations for further research.

First, operationalizing the competitiveness of markets is challenging, and private companies often control access to the necessary price data, which they tend to consider sensitive and confidential business information. This may explain why hardly any evidence is available on rent-extraction by large processing and exporting companies, and also raises some constraints to my research. While I got access to essential price data, the available data was limited to one cooperative. Chapter 5 makes a parallel trends assumption on the value of high-quality versus low-quality produce, which is untestable given the available data. Moreover, I study market power in a situation where it is most likely to arise: in a market with only few buyers, and this complicates the construction of standard errors. As I will explain below, analyzing price discrimination from processors across cooperatives could give further insight in rent-seeking, while requiring different assumptions. Such an analysis could therefore complement the analysis of Chapter 5, and could perhaps provide further confidence in the results.

Even if the results can be trusted for now, I estimated market power only over rents created by the intervention in one quality dimension, and market power over these rents might change in the future. For example, when other processors also start to demand high-quality milk, competition for high-quality milk increases, and intervention rents may be shifted back to producers, while markets for low-quality milk might completely dry up. While my research could not identify any of such market developments, the effects of the intervention are therefore less clear in the long-run.

Second, outcomes are measured by narrow indicators, like for example the adoption of food safety technology in Chapters 2 and 3, and food quality in Chapter 4, and prices in Chapters 5 and 6. As I will further discuss below, interventions may have led farmers to divert scarce mental and financial resources away from other activities, and the costs of this are not included in these narrow outcome indicators. Meanwhile, statistical power was insufficient to study effects on broad welfare indicators.

7.4.1. Competitiveness of markets

Many global food value chains are characterized by widespread market concentration, especially at the level of processors and exporters (FAO 2014, Gaji and Tsowou 2015, Grabs 2017). Chapter 5 has shown that the resulting imperfect competition may have strong implications for the formation of prices and the distribution of rents along value chains. Moreover, it may affect incentives for investments, and shape the effects of interventions that intend to help poor producers (Macchiavello and Morjaria 2017).

Still, the competitiveness of intermediate markets remains highly understudied. As the industrial organization literature typically focuses on consumer welfare, and farm gate prices are only a small fraction of consumer prices, it has given relatively little attention to effects of imperfect competition on smallholder farmers (Grau and Hockmann 2018). Meanwhile, the development economics literature has mostly focused on competition between grain traders, among which data collections and experiments might be relatively easily and cheaply organized (Bergquist 2017, Casaburi and Reed 2017, Dillon and Dambro 2017).

To address the knowledge gap, there is a need for both proper theoretical frameworks as well as rigorous empirical analysis. Theoretical models might shed light on (i) how market structures affect the formation of prices, (ii) how market structures shape the distribution of rents of interventions that increase the quality and quantity of food produced by smallholders, (iii) how market structures shape incentives for intermediaries to offer interlinked contracts, (iv) how interventions can intendedly or unintendedly change market structures, and (v) how complementary interventions can mitigate negative side-effects of changing market structures.

A thorough theoretical understanding is also needed for a proper interpretation of empirical results. This becomes clear, for example, when we compare the paper of Casaburi and Reed

(2017) with the paper of Bergquist (2017). The former randomizes premiums for cocoa among individual traders, and finds small differences in the prices paid by treated and control traders. Using a model of price differentiation, they interpret these small price differences as evidence for relatively competitive markets. The latter study randomizes cost-reductions and trader entry at the level of local markets, assumes these markets to be independent, and interprets the small differences in prices paid by traders in treated and control markets as evidence for imperfectly competitive markets. The seemingly contradicting conclusions at least suggests that interpreting results through the wrong theoretical lens can cause conclusions to reverse, and that conclusions should not be drawn too easily.

Empirical research can help understand how widespread the phenomenon of imperfectly competitive markets actually is, and how this varies with food product characteristics (e.g. perishability and the degree of processing required), the segment of the market (e.g. regular or premium quality) and the stage in the value chain (e.g. among traders or processors and exporters). Some first empirical contributions could provide simple market share estimates. While market concentration does not automatically give actors market power, it may provide some first indication where markets are more likely to be imperfectly competitive. Although market shares may be easier to estimate than market power, even market share estimates are currently often not publicly available, neither for local nor for national markets. And when market share estimates are available, they often concern seller market shares rather than buyer market shares.

Further empirical work can follow the spirit of Chapter 5, and estimate price impacts of interventions that benefit smallholders producers to learn more about the competitiveness of markets and how this affects the distribution of intervention rents. Such studies should take into account the methodological insights of this Chapter. First, since rent extraction can take place at aggregate level, prices should be studied at sufficiently aggregate level. Especially when studying market power among large processors and exporters, the proposed Difference-in-Differences is then often the best feasible strategy that one can use to identify effects on prices. Second, since rents might not be extracted directly, the evaluator should study impacts over a long period of time. Third, to study long-run price impacts, the evaluator should control for other events that influence prices and needs a credible proxy for the counterfactual. In Chapter 5 the price paid by another buyer served this role, but if buyers can discriminate prices, one can also think of prices paid by the same buyer to other producers. The latter might be more easy to collect from the private investigator participating in the project if negotiated in advance, and requires different assumptions to proxy the value of produce for the buyer. However, as this latter strategy allows the buyer to affect the counterfactual estimate, it should be very costly to affect the counterfactual estimate for the identification to be reasonably manipulation-robust. Fourth, the impact on prices can only be identified if these imports are sufficiently large.

Research along these lines has two significant limitations. First, while the assumptions needed for causal inference in the Difference-in-Difference framework may be plausible, they cannot always be empirically verified. This makes replication studies in similar settings even more important. Second, it does not estimate market power over total rents, but only over rents affected by the specific intervention. While market power over specific intervention rents may be most relevant for policy decisions on such interventions, one may want to learn more about the general competitiveness of markets for other purposes.

Since interlinked contracts are often offered by traders that at most have some very local market power (Casaburi and Reed 2017, Macchiavello and Morjaria 2017), and since there are many traders and local markets, smartly designed experiments might help to understand incentives to offer interlinked contracts to smallholder farmers. For example, by offering a premium to a proportion of traders that randomly varies across local markets, exogenous variation might be created in the effective competition experienced by traders in their very local context. This exogenous variation could then be used to study the effects on the provision and terms of interlinked contracts.

Another line of empirical research should identify how value chain interventions affect market structures. For example, one of the top priorities of donor agencies is to connect smallholder farmers to modern value chains (OECD/WTO 2017), and it has become more common to involve value chain actors in interventions to reach this objective. When such interventions benefit specific buyers, these buyers may gain market shares. Ultimately, this may lead to increased market power and rent-extraction, thus lowering positive impacts for smallholder farmers. As another example, while new information technologies can have the potential to change market structures by distributing information on market prices, returns to scale can create new monopolies among information technology service providers. Empirical research is necessary to explore to what extent interventions actually affect market structures and lead to rent-extraction.

7.4.2. Measuring welfare effects of interventions

Substitution effects

Even in the absence of rent-extraction by buyers, the welfare effects of productivity- and quality enhancing interventions have often remained unclear (Bulte et al. 2014). As argued in Chapter 2, and potentially also valid for the situations described in later Chapters, interventions may lead farmers to divert scarce mental and financial resources away from other activities to be able to invest them in new technologies. Narrowly defined outcome indicators may miss the lost productivity in these other activities, and evaluations may be underpowered to detect effects on more broadly defined outcome indicators that are closer to welfare, especially when these evaluations are based on a clustered sampling design.

Inference with a small number of clusters

Many impact evaluations use a clustered sampling design, because interventions are targeted at higher level clusters (e.g. farmer groups, producer organizations, villages or districts) with all the individuals within the clusters being treated, or because researchers want to limit spillovers between treatment and control respondents. Due to various constraints, researchers may be restricted in the selection of clusters. For example, the sampling designs used in this thesis faced constraints regarding (i) the budget for the study, (ii) the number of existing farmer groups in the area, (ii) the number of treated clusters, (iii) the number of comparable control clusters, and (iv) the availability of baseline data.

If the number of sampled clusters is low, inference based on cluster-robust standard errors leads to over-rejecting the null hypothesis, and recent studies therefore use the wild bootstrap as recommended by Cameron, Gelbach, and Miller (2008). The wild bootstrap procedure avoids over-rejection of the null hypothesis and leads to reliable statistical inference even when the number of clusters is low if there are at least some treated and untreated clusters (MacKinnon and Webb 2017, 2018).

Power calculations with a small number of clusters

While reliable inference in a cluster design is possible with a wild bootstrap procedure, the cluster design leads to decreased statistical power to distinguish an actual treatment effect from a null effect. In the next few paragraphs, I will more formally argue why many studies fail to detect welfare impacts and indicate how badly power can be affected by cluster designs.

Suppose one aims to estimate:

$$y_{ic} = \gamma + \beta T_c + u_c + v_{ic}, \qquad (7.1)$$

where y_{ic} is the outcome variable of interest, T_c is a randomized binary treatment, $u_c \sim IIDN(0, \sigma_u^2)$ is the cluster-level error term, and $v_{ic} \sim IIDN(0, \sigma_v^2)$ the individual-level error term. For ease of expression, I define $\varepsilon_{ic} \equiv u_c + v_{ic}$, so that $\sigma^2 \equiv Var(\varepsilon_{ic}) = \sigma_u^2 + \sigma_v^2$ and $\rho \equiv Corr(\varepsilon_{ic}, \varepsilon_{jc}) = \sigma_u^2/\sigma^2$ for $i \neq j$.

Then, the Minimum Detectable Effect (MDE) is given by:

$$\beta_{MDE} = (t_{\alpha/2} + t_{1-\kappa})\sigma_{\hat{\beta}}, \qquad (7.2)$$

where $t_{\alpha/2}$ and $t_{1-\kappa}$ are critical values of the t-distribution with α the significance level and κ the power, and:

$$\sigma_{\hat{\beta}}^2 \equiv Var(\hat{\beta}) = \frac{1}{\bar{T}(1-\bar{T})}\left(\frac{\sigma_u^2}{N_c} + \frac{\sigma_v^2}{N}\right) = \frac{1}{\bar{T}(1-\bar{T})}\frac{\sigma^2}{N}[1 + \rho(m-1)], \qquad (7.3)$$

where $\bar{T}$ is the proportion that received treatment, N_c is the number of clusters, N is the total number of individuals, and $m \equiv N/N_c$ is the cluster size (Duflo, Glennerster, and Kremer 2007). It is common to use the t distribution with $N_c - 1$ degrees of freedom (Cameron and Miller 2015). The latter term $1 + \rho(m-1)$ is also known as the design effect, as it captures the correction for within-cluster correlation of the error terms, and is caused by the clustered sampling design.

Substituting (7.3) in (7.2) and little rearrangement yields the following estimate for the standardized MDE:

$$\frac{\beta_{MDE}}{\sigma} = (t_{\alpha/2} + t_{1-\kappa})\sqrt{\frac{1}{\bar{T}(1-\bar{T})}}\sqrt{\frac{1}{N}}\sqrt{1 + \rho(m-1)}, \qquad (7.4)$$

Equation (7.4) shows that, when holding the sample size N constant, increasing the cluster size m may increase the standardized MDE in two ways. First, the design effect increases if observations are correlated within the cluster ($\rho > 0$). This term reflects the increased disturbance of cluster-level error term u_c. Secondly, the number of clusters $N_c = N/m$ and, thus, the degrees of freedom decrease, which causes the critical values of the t distribution to increase. To illustrate how large the standardized MDE can become, Table 7.1 shows the standardized MDE for sample sizes similar used in this thesis and varying levels of intra-cluster correlations.

Table 7.1. Standardized MDE

Chapter	N	N_c	ρ	β_{MDE}/σ
3	2713	152	0.00	0.109
3	2713	152	0.20	0.228
3	2713	152	0.40	0.303
3	2713	152	0.60	0.364
5	752	13	0.00	0.221
5	752	13	0.20	0.776
5	752	13	0.40	1.076
5	752	13	0.60	1.308
$\alpha = 0.05, \kappa = 0.8, \bar{T} = 0.5$				

While for given fixed intra-cluster correlation ρ, clustering similarly affects the standardized MDE of narrowly and broadly defined outcomes indicators, actual standardized treatment effects are likely to be lower for more broadly defined outcomes. For narrowly defined outcome indicators, like the adoption of specific technologies used in Chapter 2 and 3, the quality measures used in Chapter 4, the treatment effect β may be large compared to the variance σ^2 in these variables. Instead, even when effects on these narrowly defined outcome indicators fully translate into higher over-all investments, income or consumption, the higher variance σ^2 in these variables may cause standardized treatment effects to be too small to be detected.

Increasing power

To measure the effect on broad outcome indicators, power needs to be increased. This can be done in several ways. First, increasing the cluster size has a modest effect if outcomes are correlated within the cluster, as it does not decrease the effect of the variance of the cluster-level error term u_c. Second, proportionally increasing the number of clusters is more effective as it (i) decreases the effect of both the cluster-level error term u_c and the individual-level error term v_{ic}, and (ii) increases the degrees of freedom used to obtain $t_{\alpha/2} + t_{1-\kappa}$. Third, one may reduce the unexplained variance σ^2. Generally, the unexplained variance can be reduced by controlling for covariates that are not affected by the treatment. If the dependent variable is highly autocorrelated and precisely measured, including the baseline dependent variable in the set of covariates further decreases the unexplained variance σ^2. Instead, if the dependent variable of interest is less autocorrelated or less precisely measured, averaging this variable over multiple follow-ups and measurements can reduce the unexplained variance σ^2 (McKenzie 2012). Fourth, one may induce and exploit exogenous within-cluster variation in take-up of the intervention, and in this way circumvent the cluster design.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" *National Bureau of Economic Research Working Paper Series* No. 24003. doi: 10.3386/w24003.

Abate, Gashaw Tadesse, and Tanguy Bernard. 2017. "Farmers' Quality Assessment of Their Crops and Its Impact on Commercialization Behavior." *IFPRI Discussion Paper 1624*.

Ahmed, M. F., S. Ahuja, M. Alauddin, S. J. Hug, J. R. Lloyd, A. Pfaff, T. Pichler, C. Saltikov, M. Stute, and A. van Geen. 2006. "Ensuring Safe Drinking Water in Bangladesh." *Science* 314 (5806):1687-1688. doi: 10.1126/science.1133146.

Ashraf, Nava, Xavier Giné, and Dean Karlan. 2009. "Finding Missing Markets (and a Disturbing Epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya." *American Journal of Agricultural Economics* 91 (4):973-990. doi: 10.1111/j.1467-8276.2009.01319.x.

Bandyopadhyay, R., A. Ortega-Beltran, A. Akande, C. Mutegi, J. Atehnkeng, L. Kaptoge, A.L. Senghor, B.N. Adhikari, and P.J. Cotty. 2016. "Biological control of aflatoxins in Africa: current status and potential challenges in the face of climate change." *World Mycotoxin Journal* 9 (5):771-789. doi: 10.3920/wmj2016.2130.

Barrett, Christopher B., Maren E. Bachke, Marc F. Bellemare, Hope C. Michelson, Sudha Narayanan, and Thomas F. Walker. 2012. "Smallholder Participation in Contract Farming: Comparative Evidence from Five Countries." *World Development* 40 (4):715-730. doi: https://doi.org/10.1016/j.worlddev.2011.09.006.

Bergquist, Lauren Falcao. 2017. "Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya." *Working Paper*.

Bernard, Tanguy, Alain de Janvry, Samba Mbaye, and Elisabeth Sadoulet. 2017. "Expected Product Market Reforms and Technology Adoption by Senegalese Onion Producers." *American Journal of Agricultural Economics* 99 (4):1096-1115. doi: 10.1093/ajae/aax033.

Bernard, Tanguy, Agnes Le Port, Melissa Hidrobo, and Rahul Rawat. 2019. "Nutrition-Based Incentives in Dairy Contract Farming in Northern Senegal." *American Journal of Agricultural Economics* 101 (2):404-435. doi: 10.1093/ajae/aay036.

Bernard, Tanguy, and David J. Spielman. 2009. "Reaching the rural poor through rural producer organizations? A study of agricultural marketing cooperatives in Ethiopia." *Food Policy* 34 (1):60-69. doi: https://doi.org/10.1016/j.foodpol.2008.08.001.

Bitzer, Verena, and Pieter Glasbergen. 2015. "Business–NGO partnerships in global value chains: part of the solution or part of the problem of sustainable change?" *Current Opinion in Environmental Sustainability* 12:35-40. doi: https://doi.org/10.1016/j.cosust.2014.08.012.

Bouma, Jetske, and Ezra Berkhout. 2015. Public-Private Partnership in Development Cooperation: Potential and Pitfalls for Inclusive Green Growth. The Hague: PBL Netherlands Environmental Assessment Agency.

Bulte, Erwin, Gonne Beekman, Salvatore Di Falco, Joseph Hella, and Pan Lei. 2014. "Behavioral Responses and the Impact of New Agricultural Technologies: Evidence from a Double-blind Field Experiment in Tanzania." *American Journal of Agricultural Economics* 96 (3):813-830. doi: 10.1093/ajae/aau015.

Burchardi, Konrad B, Selim Gulesci, Benedetta Lerva, and Munshi Sulaiman. 2019. "Moral Hazard: Experimental Evidence from Tenancy Contracts*." *The Quarterly Journal of Economics* 134 (1):281-347. doi: 10.1093/qje/qjy023.

Byerlee, Derek, Alain De Janvry, Elisabeth Sadoulet, Robert Townsend, and Irina Klytchnikova. 2008. World Development Report 2008: Agriculture for Development. Washington DC: World Bank Group.

Calì, Massimiliano, and Dirk Willem te Velde. 2011. "Does Aid for Trade Really Improve Trade Performance?" *World Development* 39 (5):725-740. doi: https://doi.org/10.1016/j.worlddev.2010.09.018.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3):414-427. doi: 10.1162/rest.90.3.414.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2):317-372. doi: 10.3368/jhr.50.2.317.

Casaburi, Lorenzo, and Rocco Macchiavello. 2015. "Loyalty, Exit, and Enforcement: Evidence from a Kenya Dairy Cooperative." *American Economic Review* 105 (5):286-90. doi: 10.1257/aer.p20151076.

Casaburi, Lorenzo, and Tristan F. Reed. 2017. "Competition in Agricultural Markets: An Experimental Approach." *CEPR Discussion Paper* No. DP11985.

Center for Global Development and Social Finance. 2013. Investing in Social Outcomes: Development Impact Bonds. Washington DC/London: Center for Global Development/Social Finance.

Christiaensen, Luc, Lionel Demery, and Jesper Kuhl. 2011. "The (evolving) role of agriculture in poverty reduction—An empirical perspective." *Journal of Development Economics* 96 (2):239-254. doi: https://doi.org/10.1016/j.jdeveco.2010.10.006.

Cochrane, Willard W. 1958. *Farm Prices: Myth and Reality*. St. Paul: University of Minnesota Press.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S. Lee. 2019. "“Nash-in-Nash” Bargaining: A Microfoundation for Applied Work." *Journal of Political Economy* 127 (1):163-195. doi: 10.1086/700729.

Cotty, Peter J, Larry Antilla, and Phillip J Wakelyn. 2007. "Competitive exclusion of aflatoxin producers: Farmer-driven research and development." In *Biological control: a global perspective*, edited by Charles Vincent, Mark S. Goettel and George Lazarovits, 241-253. Oxfordshire, UK: CAB International.

D’Espallier, Bert, Jann Goedecke, Marek Hudon, and Roy Mersland. 2017. "From NGOs to Banks: Does Institutional Transformation Alter the Business Model of Microfinance Institutions?" *World Development* 89:19-33. doi: https://doi.org/10.1016/j.worlddev.2016.06.021.

Daniel, Johnni H., Lauren W. Lewis, Yanique A. Redwood, Stephanie Kieszak, Robert F. Breiman, W. Dana Flanders, Carlos Bell, John Mwihia, George Ogana, Sopiato Likimani, Masja Straetemans, and Michael A. McGeehin. 2011. "Comprehensive Assessment of Maize Aflatoxin Levels in Eastern Kenya, 2005–2007." *Environmental Health Perspectives* 119 (12):1794-1799. doi: 10.1289/ehp.1003044.

De Brauw, Alan, Berber Kramer, Hazel Malapit, Elena Martinez, and Mike Murphy. 2018. Impact Evaluation Associated with the Bangladesh AVC Project: Midline Report. Washington: IFPRI.

De La O Campos, A.P., C. Villani, B. Davis, and M. 2018 Takagi. 2018. Ending Extreme Poverty in Rural Areas: Sustaining Livelihoods to Leave No One Behind. Rome: FAO.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11):3266-3302. doi: doi: 10.1257/aer.20171330.

Dercon, Stefan, and Luc Christiaensen. 2011. "Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia." *Journal of Development Economics* 96 (2):159-173. doi: https://doi.org/10.1016/j.jdeveco.2010.08.003.

Dercon, Stefan, and Douglas Gollin. 2014. "Agriculture in African Development: Theories and Strategies." *Annual Review of Resource Economics* 6 (1):471-492. doi: 10.1146/annurev-resource-100913-012706.

Dillon, Brian, and Chelsey Dambro. 2017. "How Competitive Are Crop Markets in Sub-Saharan Africa?" *American Journal of Agricultural Economics* 99 (5):1344-1361. doi: 10.1093/ajae/aax044.

Dolan, C., and J. Humphrey. 2000. "Governance and Trade in Fresh Vegetables: The Impact of UK Supermarkets on the African Horticulture Industry." *The Journal of Development Studies* 37 (2):147-176. doi: 10.1080/713600072.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*, 3895-3962. Elsevier.

Dupas, Pascaline. 2014. "Getting essential health products to their end users: Subsidize, but how much?" *Science* 345 (6202):1279-1281. doi: 10.1126/science.1256973.

Easterly, William. 2009. "Can the West Save Africa?" *Journal of Economic Literature* 47 (2):373-447. doi: 10.1257/jel.47.2.373.

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H. Dar. 2016. "Technological Innovations, Downside Risk, and the Modernization of Agriculture." *American Economic Review* 106 (6):1537-61. doi: 10.1257/aer.20150474.

FAO. 2014. The Changing Role of Multinational Companies in the Global Banana Trade. Rome: FAO.

Farina, Elizabeth M. M. Q., Graciela E. Gutman, Pablo J. Lavarello, Rubens Nunes, and T. Reardon. 2005. "Private and Public Milk Standards in Argentina and Brazil." *Food Policy* 30 (3):302-315. doi: 10.1016/j.foodpol.2005.05.008.

Fischer, Elisabeth, and Matin Qaim. 2012. "Linking Smallholders to Markets: Determinants and Impacts of Farmer Collective Action in Kenya." *World Development* 40 (6):1255-1268. doi: https://doi.org/10.1016/j.worlddev.2011.11.018.

Foster, Andrew D., and Mark R. Rosenzweig. 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics* 2 (1):395-424. doi: 10.1146/annurev.economics.102308.124433.

Gaji, Samuel K., and Komi Tsowou. 2015. Cocoa Industry: Integrating Small Farmers into the Global Value Chain. New York and Geneva: United Nations.

Gong, Yunyun, Assomption Hounsa, Sharif Egal, Paul C. Turner, Anne E. Sutcliffe, Andrew J. Hall, Kitty Cardwell, and Christopher P. Wild. 2004. "Postweaning Exposure to Aflatoxin

Results in Impaired Child Growth: A Longitudinal Study in Benin, West Africa." *Environmental Health Perspectives* 112 (13):1334-1338. doi: 10.1289/ehp.6954.

Grabs, Janina. 2017. "The Rise of Buyer-Driven Sustainability Governance: Emerging Trends in the Global Coffee Sector." *ZenTra Working Papers in Transnational Studies* No. 73/2017.

Grau, Aaron, and Heinrich Hockmann. 2018. "Market power in the German dairy value chain." *Agribusiness* 34 (1):93-111. doi: 10.1002/agr.21529.

Grunert, Klaus G. 2005. "Food quality and safety: consumer perception and demand." *European Review of Agricultural Economics* 32 (3):369-391. doi: 10.1093/eurrag/jbi011.

Haggbladde, Steven, Peter B.R. Hazell, and Paul A. Dorosh. 2007. "Sectoral Growth Linkages Between Agriculture and the Rural Nonfarm Economy." In *Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World*, edited by Steven Haggblade, Peter B. R. Hazell and Thomas Reardon, 141-182. Baltimore: The Johns Hopkins University Press.

Hart, Oliver, and Bengt Holmström. 1987. "The theory of contracts." In *Advances in Economic Theory: Fifth World Congress*, edited by Truman Fassett Bewley, 71-156. Cambridge: Cambridge University Press.

Hoffmann, Vivian, and Kelly Jones. 2018. "Improving Food Safety on the Farm: Experimental Evidence from Kenya on Agricultural Incentives and Subsidies as Public Health Investments." *IFPRI Discussion Paper 01746*.

Hoffmann, Vivian, Kelly Jones, and Jef L. Leroy. 2018. "The impact of reducing dietary aflatoxin exposure on child linear growth: a cluster randomised controlled trial in Kenya." *BMJ Global Health* 3 (6):e000983. doi: 10.1136/bmjgh-2018-000983.

Hoffmann, Vivian, Sarah Kariuki, Janneke Pieters, and Mark Treurniet. 2017. "Market Safe Pre-Analysis Plan." *Pre-Analysis Plan*.

Hoffmann, Vivian, Sarah Kariuki, Janneke Pieters, and Mark Treurniet. 2018. "Can Markets Support Smallholder Adoption of a Food Safety Technology?" *IFPRI Project Note*.

Hoffmann, Vivian, Nicholas Magnan, Gissele Gajate Garrido, Daniel Akwasi Kanyam, and Nelson Opuku. 2018. "Information, Technology, and Market Rewards: Incentivizing Aflatoxin Control in Ghana." *Working Paper*.

Hoffmann, Vivian, and Christine Moser. 2017. "You get what you pay for: the link between price and food safety in Kenya." *Agricultural Economics* 48 (4):449-458. doi: 10.1111/agec.12346.

Hoffmann, Vivian, Christine Moser, and Timothy Herrman. 2019. "Demand for Aflatoxin-Safe Maize in Kenya: Dynamic Response to Price and Advertising." *Working Paper*.

Hoffmann, Vivian, Christine Moser, and Alexander E. Saak. 2019. "Food Safety in Low and Middle-Income Countries: The Evidence Through an Economic Lens." *Working Paper*.

Hoffmann, Vivian, Samuel Mutiga, Jagger Harvey, Rebecca Nelson, and Michael Milgroom. 2013. "Asymmetric Information and Food Safety: Maize in Kenya." *Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington DC*.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20 (1):1-24. doi: DOI: 10.1093/pan/mpr013.

Indonesian Ministry of Trade. 2019a. "Growth of Non-Oil and Gas Export (Commodity) Period: 2013-2018." Accessed January 21, 2019. http://www.kemendag.go.id/en/economic-profile/indonesia-export-import/growth-of-non-oil-and-gas-export-commodity.

Indonesian Ministry of Trade. 2019b. "Growth of Non-Oil and Gas Import (Commodity) Period: 2013-2018." Accessed January 21, 2019. http://www.kemendag.go.id/en/economic-profile/indonesia-export-import/growth-of-non-oil-and-gas-import-commodity.

Jack, B. Kelsey. 2013. Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries. PAL (MIT) and CEGA (UC Berkeley): Agricultural Technology Adoption Initiative.

Just, Richard E., and Rulon D. Pope. 1978. "Stochastic specification of production functions and economic implications." *Journal of Econometrics* 7 (1):67-86. doi: https://doi.org/10.1016/0304-4076(78)90006-4.

Leroy, Jef L., Jia-Sheng Wang, and Kelly Jones. 2015. "Serum aflatoxin B1-lysine adduct level in adult women from Eastern Province in Kenya depends on household socio-economic status: A cross sectional study." *Social Science & Medicine* 146:104-110. doi: https://doi.org/10.1016/j.socscimed.2015.10.039.

Ligon, Ethan, and Elisabeth Sadoulet. 2007. "Estimating the Effects of Aggregate Agricultural Growth on the Distribution of Expenditures." *Background Paper for the World Development Report 2008*.

Macchiavello, Rocco, and Ameet Morjaria. 2017. "Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills." *Buffett Institute Global Poverty Research Lab Working Paper No. 17-103*.

R

MacKinnon, James G., and Matthew D. Webb. 2017. "Wild Bootstrap Inference for Wildly Different Cluster Sizes." *Journal of Applied Econometrics* 32 (2):233-254. doi: 10.1002/jae.2508.

MacKinnon, James G., and Matthew D. Webb. 2018. "The wild bootstrap for few (treated) clusters." *The Econometrics Journal* 21 (2):114-135. doi: 10.1111/ectj.12107.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149):976-980. doi: 10.1126/science.1238041.

Massa, Isabella. 2013. "Impact of Aid for Trade: A Survey of Empirical Evidence." In *Assessing Aid for Trade: Effectiveness, Current Issues and Future Directions*, edited by Mohammad A. Razzaque and Dirk Willem te Velde, 14-32. London: Commonwealth Secretariat.

McCarney, Rob, James Warner, Steve Iliffe, Robbert van Haselen, Mark Griffin, and Peter Fisher. 2007. "The Hawthorne Effect: a Randomised, Controlled Trial." *BMC Medical Research Methodology* 7 (1):30. doi: 10.1186/1471-2288-7-30.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99 (2):210-221. doi: https://doi.org/10.1016/j.jdeveco.2012.01.002.

Mia, Md Aslam, and Hwok-Aun Lee. 2017. "Mission drift and ethical crisis in microfinance institutions: What matters?" *Journal of Cleaner Production* 164:102-114. doi: https://doi.org/10.1016/j.jclepro.2017.06.176.

Morey, Phillip. 2011. Dairy Industry Development in Indonesia. Jakarta: International Finance Corporation.

Munasib, Abdul, and Devesh Roy. 2011. "Sanitary and Phytosanitary Standards as Bridge to Cross." *IFPRI Discussion Paper 01140*.

Mutiga, S. K., V. Were, V. Hoffmann, J. W. Harvey, M. G. Milgroom, and R. J. Nelson. 2014. "Extent and Drivers of Mycotoxin Contamination: Inferences from a Survey of Kenyan Maize Mills." *Phytopathology* 104 (11):1221-1231. doi: 10.1094/phyto-01-14-0006-r.

Muto, Megumi, and Takashi Yamano. 2009. "The Impact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda." *World development* 37 (12):1887-1896. doi: 10.1016/j.worlddev.2009.05.004.

Newbery, David M. G., and Joseph E. Stiglitz. 1984. "Pareto Inferior Trade." *The Review of Economic Studies* 51 (1):1-12. doi: 10.2307/2297701.

OECD. 2018. Development Co-operation Report 2018: Joining Forces to Leave No One Behind. Paris: OECD Publishing.

OECD/WTO. 2017. Aid for Trade at a Glance 2017: Promoting Trade, Inclusiveness and Connectivity for Sustainable Development. Geneva/Paris: WTO/OECD Publishing.

Poulton, Colin, and Jon Macartney. 2012. "Can Public–Private Partnerships Leverage Private Investment in Agricultural Value Chains in Africa? A Preliminary Review." *World Development* 40 (1):96-109. doi: https://doi.org/10.1016/j.worlddev.2011.05.017.

Ravallion, Martin. 2012. "Fighting Poverty One Experiment at a Time: Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty: Review Essay." *Journal of Economic Literature* 50 (1):103-14. doi: 10.1257/jel.50.1.103.

Razzaque, Mohammad A., and Dirk Willem te Velde. 2013. "Aid for Trade: Effectiveness, Current Issues and Future Directions: An Overview." In *Assessing Aid for Trade: Effectiveness, Current Issues and Future Directions*, edited by Mohammad A. Razzaque and Dirk Willem te Velde, 1-13. London: Commonwealth Secretariat.

Reardon, Thomas, Christopher B. Barrett, Julio A. Berdegué, and Johan F. M. Swinnen. 2009. "Agrifood Industry Transformation and Small Farmers in Developing Countries." *World Development* 37 (11):1717-1727. doi: 10.1016/j.worlddev.2008.08.023.

Rodrigues, Angela M., Nicola O'Brien, David P. French, Liz Glidewell, and Falko F. Sniehotta. 2015. "The Question–Behavior Effect: Genuine Effect or Spurious Phenomenon? A Systematic Review of Randomized Controlled Trials with Meta-Analyses." *Health Psychology* 34 (1):61-78. doi: 10.1037/hea0000104.

Rosenzweig, Mark R., and Hans P. C. Binswanger. 1993. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *The Economic Journal* 103 (416):56-78. doi: 10.2307/2234337.

Saenger, Christoph, Matin Qaim, Maximo Torero, and Angelino Viceisza. 2013. "Contract Farming and Smallholder Incentives to Produce High Quality: Experimental Evidence from the Vietnamese Dairy Sector." *Agricultural Economics* 44 (3):297-308. doi: 10.1111/agec.12012.

Saenger, Christoph, Maximo Torero, and Matin Qaim. 2014. "Impact of Third-party Contract Enforcement in Agricultural Markets—A Field Experiment in Vietnam." *American Journal of Agricultural Economics* 96 (4):1220-1238. doi: 10.1093/ajae/aau021.

Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan. 2016. "The Psychological Lives of the Poor." *American Economic Review* 106 (5):435-40. doi: 10.1257/aer.p20161101.

Schreiner, Mark. 2012. A Simple Poverty Scorecard for Indonesia. Progress out of Poverty.

Sexton, Richard J. 2012. "Market Power, Misconceptions, and Modern Agricultural Markets." *American Journal of Agricultural Economics* 95 (2):209-219. doi: 10.1093/ajae/aas102.

Shipe, W. F. 1959. "The Freezing Point of Milk. A Review." *Journal of Dairy Science* 42 (11):1745-1762. doi: 10.3168/jds.S0022-0302(59)90805-7.

Smits, Jeroen, and Roel Steendijk. 2015. "The International Wealth Index (IWI)." *Social Indicators Research* 122 (1):65-85. doi: 10.1007/s11205-014-0683-x.

Stiglitz, Joseph E., and Andrew Charlton. 2013. "The Right to Trade: Rethinking the Aid for Trade Agenda." In *Assessing Aid for Trade: Effectiveness, Current Issues and Future Directions*, edited by Mohammad A. Razzaque and Dirk Willem te Velde, 359-397. London: Commonwealth Secretariat.

Strosnider, Heather, Eduardo Azziz-Baumgartner, Marianne Banziger, Ramesh V. Bhat, Robert Breiman, Marie-Noel Brune, Kevin DeCock, Abby Dilley, John Groopman, Kerstin Hell, Sara H. Henry, Daniel Jeffers, Curtis Jolly, Pauline Jolly, Gilbert N. Kibata, Lauren Lewis, Xiumei Liu, George Luber, Leslie McCoy, Patience Mensah, Marina Miraglia, Ambrose Misore, Henry Njapau, Choon-Nam Ong, Mary T.K. Onsongo, Samuel W. Page, Douglas Park, Manish Patel, Timothy Phillips, Maya Pineiro, Jenny Pronczuk, Helen Schurz Rogers, Carol Rubin, Myrna Sabino, Arthur Schaafsma, Gordon Shephard, Joerg Stroka, Christopher Wild, Jonathan T. Williams, and David Wilson. 2006. "Workgroup Report: Public Health Strategies for Reducing Aflatoxin Exposure in Developing Countries." *Environmental Health Perspectives* 114 (12):1898-1903. doi: 10.1289/ehp.9302.

Swinnen, Johan F. M., and Miet Maertens. 2007. "Globalization, Privatization, and Vertical Coordination in Food Value Chains in Developing and Transition Countries." *Agricultural Economics* 37 (s1):89-102. doi: 10.1111/j.1574-0862.2007.00237.x.

Swinnen, Johan, and Anneleen Vandeplas. 2011. "Rich Consumers and Poor Producers: Quality and Rent Distribution in Global Value Chains." *Journal of Globalization and Development* 2 (2):1-30. doi: 10.1515/1948-1837.1036.

Turner, Paul C., Andrew C. Collinson, Yin Bun Cheung, YunYun Gong, Andrew J. Hall, Andrew M. Prentice, and Christopher P. Wild. 2007. "Aflatoxin exposure in utero causes growth faltering in Gambian infants." *International Journal of Epidemiology* 36 (5):1119-1125. doi: 10.1093/ije/dym122.

Van Beuningen, Coen, and Peter Knorringa. 2009. *Inclusive Improvement: Standards And Smallholders: Taking Stock and Moving On*. The Hague: HIVOS/ISS.

Vijil, Mariana, and Laurent Wagner. 2012. "Does Aid for Trade Enhance Export Performance? Investigating the Infrastructure Channel." *The World Economy* 35 (7):838-868. doi: 10.1111/j.1467-9701.2012.01437.x.

WHO. 2015. WHO Estimates of the Global Burden of Foodborne Diseases: Foodborne Diseases Burden Epidemiology Reference Group 2007-2015. Geneva: WHO.

Wijers, Gea D. M. 2019. "Inequality Regimes in Indonesian Dairy Cooperatives: Understanding Institutional Barriers to Gender Equality." *Agriculture and Human Values*. doi: 10.1007/s10460-018-09908-9.

World Bank. 2019a. PovcalNet dataset.

World Bank. 2019b. "What Are Public Private Partnerships?" Accessed June 27. https://ppp.worldbank.org/public-private-partnership/overview/what-are-public-private-partnerships.

Wouters, A.P. 2009. Dairy Sector Development Indonesia: Options for cooperation with The Netherlands. Wageningen, the Netherlands: Wageningen UR Livestock Research.

WTO. 2019. "Aid for Trade fact sheet." Accessed January 21, 2019. https://www.wto.org/english/tratop_e/devel_e/a4t_e/a4t_factsheet_e.htm.

WTO Hong Kong Ministerial Declaration. 2005. "Ministerial Declaration." Accessed July 1. https://www.wto.org/english/thewto_e/minist_e/min05_e/final_text_e.htm.

Zwane, Alix Peterson, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, Florencia Devoto, Bruno Crepon, and Abhijit Banerjee. 2011. "Being Surveyed Can Change Later Behavior and Related Parameter Estimates." *Proceedings of the National Academy of Sciences of the United States of America* 108 (5):1821-1826. doi: 10.1073/pnas.1000776108.

Summary

Since the launch of the 2005 Aid for Trade initiative, it has become more popular to disperse development assistance via public-private partnerships (PPPs). Yet, research on the impact of aid dispersed via such PPPs has been limited. This thesis studies the (potential) contribution of PPPs to increased food safety (SDG 2) and increased farmer incomes (SDG 1) in lower middle-income countries.

In Chapter 1, the General Introduction, I introduce the topic and main research question, and provide an overview of the methodologies used throughout this thesis.

Chapter 2 studies the impact of being surveyed, and find that this increases the adoption of food safety technology among Kenyan subsistence farmers, thus suggesting that focusing attention to the issue of food safety is an important driver of food safety improvement.

Chapter 3 and 4 study how introducing price incentives can help to increase food safety. Chapter 3 finds that providing a modest market premium for safe maize helps Kenyan subsistence farmers to increase the intensity of adoption of food safety technology. In poor years, the resulting safe maize could be consumed at home. In good years, the household would receive a small compensation for the safe maize delivered to the market.

Chapter 4 studies how a PPP intervention in the Indonesian dairy value chain increased the quality of milk delivered by members of a large dairy cooperative. Part of the positive impacts on the hygienic and compositional quality of milk are explained by the introduction of an individual price incentive system (that replaced a group-based system).

Chapters 5 and 6 study the distribution of rents created by this PPP. Chapter 5 studies how intervention rents are vertically distributed between the dairy processor and the cooperative that represented the smallholder producers, and finds that farmers received higher prices in the short run, but the processor captured the full quality rents after the intervention was completed.

Chapter 6 studies the horizontal distribution of benefits between farmers, and finds that the PPP intervention reached mainly large farmer groups and wealthy farmers with relatively many cows. While this may be the most effective way to affect large amounts of milk, the intervention helped these included farmers to earn higher prices, and therefore increased economic inequality within the cooperative membership. Since farmers did not benefit from higher prices

on average, this suggests that the poorer subsample may have impoverished relative to the counterfactual without the intervention.

Chapter 7, the General Discussion, synthesizes research findings in the light of existing literature, and argues that that the alignment of private interests with development goals is key to realize the intended development outcomes. Especially for the realization of food safety (SDG 2), private interests may be well-aligned with development goals, and PPPs may be an effective organizational tool to spur development. Instead, if the goal is to increase the income of smallholders that deliver their produce to buyers with significant market power, or the goal is to reach the poorest of the poor (SDG 1), private interests and development goals may be misaligned, and intended outcomes are less likely to be reached.

While this thesis raises critical questions to the increasingly popular trend to use PPPs to spur development, Chapter 7 further discusses some recommendations for the use of PPPs in food value chains regarding the type of activities funded, the selection of implementing partners, the incentive structure of PPPs and the creation of opportunities for learning. Chapter 7 finally highlights avenues for further research.

Samenvatting

Sinds de lancering van het 2005 Hulp voor Handel initiatief, is het steeds populairder geworden om ontwikkelingshulp uit te zetten via publiek-private partnerschappen (PPP's). Onderzoek naar de impact van hulp uitgezet via dergelijke PPP's is echter nog beperkt. Dit proefschrift bestudeert de (potentiële) bijdrage van PPP's aan verbeterde voedselveiligheid (SDG 2) en verbeterde boereninkomens (SDG 1) in lageiddeninkomenslanden.

In Hoofdstuk 1, de Algemene Introductie, introduceer ik het onderwerp en de belangrijkste onderzoeksvragen, en geef ik een overzicht van de in dit proefschrift gebruikte methodologieën.

Hoofdstuk 2 bestudeert de impact van geënquêteerd worden, en vindt dat dit de adoptie van voedselveiligheidstechnologie onder Keniaanse zelfvoorzienende boeren laat toenemen, wat suggereert dat aandacht richten op het voedselveiligheidsprobleem een belangrijke aandrijver is van voedselveiligheidsverbetering.

Hoofdstuk 3 en 4 bestuderen hoe het introduceren van prijsprikkels kan helpen om voedselveiligheid te verbeteren. Hoofdstuk 3 vindt dat het bieden van een bescheiden marktpremie voor veilige maïs Keniaanse zelfvoorzienende boeren helpt om de intensiteit van de adoptie van voedselveiligheidstechnologie te vergroten. In slechte jaren kan de resulterende veilige maïs thuis geconsumeerd worden. In goede jaren zou het huishouden een kleine compensatie ontvangen voor de veilige maïs die aan de markt geleverd wordt.

Hoofdstuk 4 bestudeert hoe een PPP interventie in de Indonesische zuivelwaardeketen de kwaliteit verbetert van melk die geleverd wordt door leden van een grote zuivelcoöperatie. Een deel van de positieve impacts op de hygiënische kwaliteit en de samenstelling van melk kan worden verklaard door de introductie van een individueel prijsprikkelsysteem (dat een systeem gebaseerd op groepen verving).

Hoofdstuk 5 en 6 bestuderen de verdeling van waarde die gecreëerd wordt door dit PPP. Hoofdstuk 5 bestudeert hoe interventiewinsten verticaal verdeeld worden tussen de zuivelverwerker en de coöperatie die de kleine boeren vertegenwoordigt, en vindt dat boeren op de korte termijn hogere prijzen ontvingen, maar dat de verwerker de volledige kwaliteitswinsten pakte nadat de interventie voltooid was.

Hoofdstuk 6 bestudeert de horizontale verdeling van voordelen tussen boeren, en vindt dat de PPP interventie met name grote boeren groepen en welvarende boeren met relatief veel koeien bereikte. Terwijl dit kan de meest effectieve manier zijn om invloed te hebben op grote hoeveelheden melk, hielp de interventie deze bereikte boeren hogere prijzen te verdienen, en heeft het zodoende de economische ongelijkheid binnen de leden van de coöperatie vergroot. Omdat boeren gemiddeld gezien niet profiteren van hogere prijzen, suggereert dit dat het armere deel van de steekproef mogelijk verarmde ten opzichte van een scenario zonder de interventie.

Hoofdstuk 7, de Algemene Discussie, presenteert een synthese van de onderzoeksbevindingen in het licht van bestaande literatuur, en betoogt dat de afstemming van private belangen met ontwikkelingsdoelen van essentieel belang is voor het realiseren van beoogde ontwikkelingsuitkomsten. In het bijzonder voor de realisatie van voedselveiligheid (SDG 2), kunnen private belangen goed afgestemd zijn met ontwikkelingsdoelen, en kunnen PPP's een effectief organisatorisch hulpmiddel zijn om ontwikkeling aan te sporen. Als echter het doel is om inkomens te verbeteren van kleine boeren die leveren aan afnemers met significante marktmacht, of het doel is om de armsten van de armsten te bereiken (SDG 1), dan zijn private belangen en ontwikkelingsdoelen mogelijk niet goed met elkaar afgestemd, en is de kans kleiner dat beoogde uitkomsten gerealiseerd worden.

Waar dit proefschrift kritische vragen stelt bij de toenemend populaire trend om PPP's te gebruiken om ontwikkeling aan te sporen, bespreekt Hoofdstuk 7 enkele aanbevelingen voor het gebruik van PPP's in voedselwaardeketens met betrekking tot het type activiteiten dat gefinancierd wordt, de selectie van implementerende partners, de beloningsstructuur van PPP's en de creatie van mogelijkheden om te leren. Hoofdstuk 7 belicht ten slotte richtingen voor toekomstig onderzoek.

Acknowledgements

This thesis would not have been possible without the special help of many people. In the coming paragraphs, I want to thank some of them in particular.

I would not have wanted to have another supervisor than Erwin Bulte. Good leaders make others perform better, and Erwin did so with me. He put trust in me from the beginning, and provided maximal freedom to both develop my thesis in my own way and intensively explore the broader field of development economics. Meanwhile, Erwin was always willing to discuss my ideas, and despite his full agenda, he always made the time to do so. Erwin's way of combining expertise with a personal aversion to hierarchy is a great example, and I am looking forward to our continuing collaboration.

My co-supervisor Jos Bijman introduced me to the world of interdisciplinary research in general, and agricultural cooperatives specifically. Jos was easily approachable, happy to discuss progress at any time in a direct and constructive way, and ready to accommodate to the planning of my research, which implied that all of my Chapters could be finished only in the last year if not the last few months.

I had the luck to work with Vivian Hoffmann, Sarah Kariuki and Janneke Pieters on the experiment described in Chapter 3. I think our work has been an excellent group effort, I liked the fruitful discussions, and being part of the team gave me a lot of energy. Sarah also helped me to correctly match the data used in Chapter 2.

Marlene Roefs was responsible for the monitoring and evaluation of activities of the public-private partnership studied in Chapter 4-6. She initially identified the opportunity to study the impact of the intervention, designed the baseline survey, connected me to key people in Indonesia, and continued to help me requesting new data. Through the past years, working with Marlene showed me her special drive to make an impact with research in difficult environments. I also want to thank Marlene's former colleague Anja Wolsky who also contributed to the survey design and supervised baseline data collection in Indonesia.

Andre Daud played a crucial role in my research in Indonesia with his excellent relations with dairy cooperatives and farmers. During my visits to local partners, he not only translated, but where necessary also further explained what I tried to say. Thanks to Andre's intermediating role, and many follow-ups, I got access to the endline data studied in Chapters 4 and 6, and even the price data studied in Chapter 5. Together with him, I want to thank his assistant Safitri Tejaningsih and their team of enumerators, including Oji, Rio and Rizky, which have done a great job by conducting thousands of surveys. Spending time with you in and around Bandung has been a great experience.

Joseph Chege and his colleagues from ACRE Africa helped me to further understand the process of the development of rainfall index insurance products. Managers of the dairy processing company and the cooperative discussed in Chapters 4-6 helped me to better to understand the local context and provided access to data.

Thousands of farmers in Kenya and Indonesia were willing to participate in surveys and made a very crucial contribution in this way.

Teun van Dommelen, Lise Masselus and Jan-Coen van der Lee provided great assistance with data entry, data cleaning and preliminary analysis.

During my past years, I enjoyed being part of the Development Economics Group of Wageningen University, a fantastic group of people. Everybody has its own inspiring motivation to jointly learn about the process of economic development, and contribute to a world without poverty. Doors have always been open for invitations for coffee, everyone was available to discuss ideas on research, teaching and outreach, and together we enjoyed various sports tournament and social gatherings. Being part of this multicultural community has truly enriched my life.

Wageningen researchers Kasper Hettinga, Simon Oosting, Gea Wijers and Bram Wouters have helped me to better understand the dynamics of milk production in developing countries.

I am also thankful for many scholars that made the time to meet me in Wageningen, Manila, Jakarta, Bandung, Norwich, Cargèse, Nairobi, Kigali, Oxford, London, Clermont-Ferrand and Washington DC, and in particular Stefan Dercon and Rob Vos, who hosted me at Oxford University and the International Food Policy Research Institute (IFPRI), respectively.

Many other people have played a supporting role, from IT service providers, caterers and cleaners at the university, to train personnel in the Netherlands, and flight attendants, taxi drivers and hotel staff abroad. I want to mention explicitly Dineke Wemmenhove, Betty de Haan, Gré Schurink and Marian Jonker, who helped to overcome bureaucratic and administrative barriers at the university.

A special thanks goes to Jonathan and Saakje Mudde, and Daniel and Hilda Garcia, who generously offered me the opportunity to stay over in Wageningen, time and time again. Your hospitality and kindness have been a very special blessing.

In the past years, I also benefitted from many opportunities earlier in my life, which have been abundant. These include the opportunities given during my studies at Tinbergen Institute and Erasmus University Rotterdam, and I want to thank especially Josse Delfgaauw, Jurjen Kamphorst, Hans van Kippersluis, Otto Swank and Bauke Visser for this. Given the vast global inequality, many people do not have the opportunities that I have had, and that eventually allowed my efforts to be translate into this thesis. I hope and pray that my work will contribute to better outcomes for some of the people that work very hard every day in circumstances that are much more challenging than mine, and yet do not see many returns.

Some of my best friends have inspired me to do what I do. My parents have always loved each other, and all four of their children, and encouraged us to use our talents. Esther, I love you and enjoy life more with you.

Above all I want to thank my heavenly Father for the hope, inspiration and guidance he provided during my PhD studies.

Mark Treurniet
Wageningen School of Social Sciences (WASS)
Completed Training and Supervision Plan

Wageningen School of Social Sciences

Name of the learning activity	Department/Institute	Year	ECTS*
A) Project related competences			
Executive Education Course on Evaluation Social Programs	J-PAL and IPA, Manilla	2015	1.4
Writing of the PhD Research Proposal	WUR	2015-2016	6
Rural Economic Analysis (AEP-31306; partly)	WUR	2016	3
Cooperatives and Producer Organisations (BEC-53306; partly)	WUR	2016	1.5
Multidisciplinary Perspectives on Quality Improvement in Value Chains	WASS and LIQUID	2016	1
Central Themes in Development Economics (DEC-30306)	WUR	2016	6
Workshop Experimental Development Economics Lab in the Field	UEA, Norwich	2016	0.3
Institutional and Organizational Economics Academy	IOEA, Cargèse	2016	1.4
B) General research related competences			
Introduction Course	WASS	2015	1
Brain Training	WGS	2015	0.3
Information Literacy Including EndNote Introduction	WGS	2016	0.6
Protecting Human Research Participants	NIH	2016	0.1
Organization of EUDN PhD Workshop	EUDN	2017	3
Writing a Book Review for a Magazine	WUR	2018	1
Scientific Writing	WGS	2018	1.8
'Does Aid for Trade Benefit Poor Producers? A Case Study'	EUDN PhD Workshop, Clermont-Ferrand	2018	1
'The competitiveness of agricultural markets: Evidence from an Aid for Trade intervention in the Indonesian dairy value chain'	CSAE Conference, Oxford	2019	1
'Aid for Trade in food value chains: Evidence on quality outcomes, price impacts and targeting from the Indonesian dairy value chain'	IFPRI, Washington DC	2019	1
C) Career related competences/personal development			
Teaching Experience and University Teaching Qualification Courses	WUR	2015-2019	4
Total			**35.4**

*One credit according to ECTS is on average equivalent to 28 hours of study load

The research described in this thesis was financially supported by the Netherlands Organisation for Scientific Research (NWO) through the Food & Business Global Challenges Programme, and the International Food Policy Research Institute (IFPRI) through the CGIAR Research Program on Policies, Institutions and Markets.